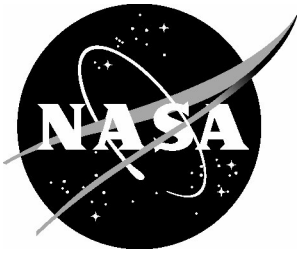
NASA/CP-2003-212642



Second Workshop on the Investigation and Reporting of Incidents and Accidents, IRIA 2003

Compiled by
Kelly J. Hayhurst and C. Michael Holloway
Langley Research Center, Hampton, Virginia

September 2003

The NASA STI Program Office . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results ... even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at ***http://www.sti.nasa.gov***

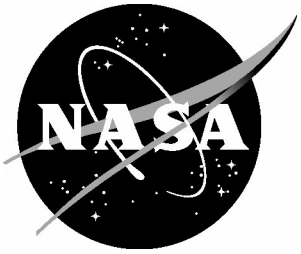
- E-mail your question via the Internet to help@sti.nasa.gov

- Fax your question to the NASA STI Help Desk at (301) 621-0134

- Phone the NASA STI Help Desk at (301) 621-0390

- Write to:
  NASA STI Help Desk
  NASA Center for AeroSpace Information
  7121 Standard Drive
  Hanover, MD 21076-1320

NASA/CP-2003-212642



Second Workshop on the Investigation and Reporting of Incidents and Accidents, IRIA 2003

Compiled by
Kelly J. Hayhurst and C. Michael Holloway
Langley Research Center, Hampton, Virginia

Proceedings of a workshop sponsored by
the National Aeronautics and Space Administration, Washington, D.C.,
and the University of Virginia, Charlottesville, Virginia,
and held at the Radisson Fort Magruder Hotel, Williamsburg, Virginia
September 16-19, 2003

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

September 2003

Available from:

NASA Center for AeroSpace Information (CASI)
7121 Standard Drive
Hanover, MD 21076-1320
(301) 621-0390

National Technical Information Service (NTIS)
5285 Port Royal Road
Springfield, VA 22161-2171
(703) 605-6000

General Chairman's Message

On behalf on NASA Langley Research Center, I welcome you to the 2nd Workshop on the Investigation and Reporting of Incidents and Accidents (IRIA03). We are pleased to be able to host this year's workshop in Williamsburg, Virginia. I hope that the workshop will be interesting and useful to you professionally, and that your stay here in Williamsburg will be enjoyable for you personally.

Putting together a workshop, even a fairly small one like IRIA, is impossible without many people working together. To the following people, I offer my special thanks for their help: Barry Strauch, Program Committee Chairman; Kelly Hayhurst, who, among many other things, put together this proceedings; Lisa Peckham, who, also among many other things, handled registrations and other logistical matters; Ray Meyer, who designed the workshop logo and printed materials; and John Knight & Kimberly Wasson, who were involved from the very beginning in all aspects of workshop planning. If you get the chance this week, please offer your thanks to these people, too.

Once again, welcome! I look forward to meeting you during the workshop. Please let me know if there is anything that I can do to help you while you're here.

C. Michael Holloway, IRIA 2003 Chairman
<http://shemesh.larc.nasa.gov/iria2003/>
NASA Langley Research Center
MS 130 / 100 NASA Road
Hampton VA 23681-2199
c.m.holloway@nasa.gov

Program Chairman's Message

On behalf of the IRIA 2003 Program Committee, I would like to welcome all of you to this year's IRIA Workshop. As last year, we are fortunate to have received many excellent submissions. The Committee, consisting of 19 specialists, represents the diversity of expertise in the field, from Europe and North America, from academia, industry, and government, and from the behavioral and computer sciences, medicine, aviation, and engineering. Each submission was reviewed by at least three Committee members, and we believe that the process has led to the high quality of papers that will be presented.

With the Workshop moving from Europe in 2002 to North America in 2003, the representation of the papers has moved as well from a majority European perspective to a majority North American one. We believe that while each Workshop gave participants a view of developments in incident and accident reporting and investigation methodology, the majority North American perspective of this year's meeting provides an enlightening contrast with last year's.

I would like to thank the members of the Program Committee for their contributions to the Workshop. As many of us, they gave of their often-busy schedules to assist in assuring a high quality Workshop. In addition, special thanks are due to the General Chair for his guidance to the Program Committee and for his overall stewardship of the Workshop.

I hope that you find the Workshop informative and rewarding.

Barry Strauch, IRIA Program Chairman
National Transportation Safety Board
490 L'Enfant Plaza SW
Washington, DC 20594
straucb@ntsb.gov

IRIA 2003 Organization

Workshop General Chairman
C. Michael Holloway, NASA Langley Research Center

Program Chairman
Barry Strauch, National Transportation Safety Board

Program Committee
S. Bogner, Institute for Study of Medical Error
D. Brown, University of Virginia
D. Busse, Microsoft
E. Byrne, National Transportation Safety Board
F. Chandler, NASA
J. Davies, University of Calgary
K. Hanks, University of Virginia
M. Holloway, NASA
C. Johnson, University of Glasgow
J. Knight, University of Virginia
P. Ladkin, University of Bielfeld
N. Leveson, M.I.T.
R. Mumaw, The Boeing Company
K. Mahoney, Thomas Jefferson National Accelerator Facility
M. O'Leary, British Airways
T. Panontin, NASA
J. Stoop, Tech. University of Delft
B. Strauch, National Transportation Safety Board
T. van der Schaaf, Technical University of Eindhoven

Sponsoring Organizations
IRIA 2003 is co-sponsored by NASA Langley Research Center and the University of Virginia.

Table of Contents

Session 7: Investigations

An electronic version of this proceedings is available on the IRIA 2003 web site:
<http://shemesh.larc.nasa.gov/iria2003/>

Archetypes for Organisational Safety

Karen Marais and Nancy G. Leveson; MIT Department of Aeronautics and Astronautics;
Cambridge, Massachusetts, U.S.A.

Keywords: organisational safety, system dynamics, archetypes

Abstract

We propose a framework using system dynamics to model the dynamic behaviour of organisations in accident analysis. Most current accident analysis techniques are event-based and do not adequately capture the dynamic complexity and non-linear interactions that characterize accidents in complex systems. In this paper we propose a set of system safety archetypes that model common safety culture flaws in organizations, i.e., the dynamic behaviour of organizations that often leads to accidents. As accident analysis and investigation tools, the archetypes can be used to develop dynamic models that describe the systemic and organizational factors contributing to the accident. The archetypes help clarify why safety-related decisions do not always result in the desired behaviour, and how independent decisions in different parts of the organisation can combine to impact safety.

Introduction

Modern socio-technical systems are becoming more complex and tightly coupled in response to increasing performance and cost requirements. Understanding these systems and analysing or accurately predicting their behaviour is often difficult. We are seeing a growing number of normal, or system, accidents, which are caused by dysfunctional interactions between components, rather than component failures. Such accidents are particularly difficult to predict or analyse [1]. Accident models focusing on direct relationships among component failure events or human errors are unable to capture these accident mechanisms adequately.

Systems and organizations continually experience change as adaptations are made in response to local pressures and short-term performance goals (e.g. productivity and cost). People adapt to their environment or they change their environment to better suit their purposes. Several decision makers at different times, in different parts of the company or organization, all striving locally to optimise performance may be preparing the stage for an accident, as illustrated by the 1987 Zeebrugge ferry disaster [2] and the Black Hawk friendly fire accident [3]. Safety defences therefore tend to degenerate systematically over time. When a larger view is taken, most accidents in complex systems can be seen to result from a migration to states of increasing risk over time. Once a system has migrated to an unsafe state, accidents are inevitable unless appropriate efforts are made to bring the system to a safe state. The Bhopal accident is a classic example.

One of the worst industrial accidents in history occurred in December 1984 at the Union Carbide chemical plant in Bhopal, India [4]. The accidental release of methyl isocyanate (MIC) resulted in at least 2000 fatalities, 10 000 permanent disabilities (including blindness), and 200 000 injuries. The Indian government blamed the accident on human error in the form of improperly performed maintenance activities. Using event-based accident models, numerous additional factors involved in the accident can be identified. But such models miss the fact that the plant had been moving over a period of many years toward a state of high-risk where almost any change in usual behaviour could lead to an accident.

If we wish to better understand past accidents and prevent future accidents, we need to look at how systems migrate towards states of increasing risk. Such understanding requires taking a long-term dynamic view of the system, and not just considering the proximate events, i.e., those events immediately preceding the actual loss event. System dynamics modelling is one way to describe dynamic change in systems. We have found it useful in understanding accidents, as argued in a companion paper, where we demonstrate its use in understanding the Walkerton *E. coli* outbreak [5]. But building system dynamics models is difficult for non-experts and usually achieved in an *ad hoc* and time-consuming manner. In developing the Walkerton system dynamics model, we found that a lot of time was needed to identify the variables of interest and to determine which relations between these variables were relevant to the accident. One way to accelerate and focus the modelling process is to start by applying archetypes that describe typical behaviour and flaws in the safety culture that have often been involved in accidents.

In this paper we propose a preliminary set of safety archetypes. The safety culture of an organization can be usefully described in terms of these safety archetypes. In accident analysis the archetypes can be used to identify and highlight change processes and the flawed decision-making that allowed the system to migrate towards an accident state. The archetypes will also form part of a new risk assessment method under development by the authors, where they will be used both as diagnostic and as prospective tools. As diagnostic or analytic tools they can be used to identify the structures underlying undesired behaviour. As synthesis tools they can be used to examine the potential undesired consequences of decisions.

This paper is organised as follows: We begin with a brief overview of system dynamics and its building blocks. Next, we propose a preliminary set of safety archetypes. In each case, illustrative examples of the archetype's application to safety are given.

System Dynamics

System dynamics is an approach to identifying, explaining, and eliminating problem behaviours in socio-economic systems, primarily by identifying feedback loops in the system. It provides a framework for dealing with dynamic complexity, where cause and effect are not obviously related. System dynamics is grounded in the theory of non-linear dynamics and feedback control, but also draws on cognitive and social psychology, organisation theory, economics, and other social sciences. For an extensive discussion, see [6].

Organisational Behaviour and the System Archetypes: System dynamics posits that the behaviour of a system arises from its structure. The structure is described in terms of feedback (causal) loops, stocks (levels) and flows (rates), and non-linearities created by interactions between system elements. In the system dynamics view, all dynamics (behaviour over time) can be explained by the interaction of the two basic types of feedback loops: positive and negative. Positive feedback loops are self-reinforcing, and are therefore referred to as *reinforcing* loops. Negative feedback loops tend to counteract change, and are therefore referred to as *balancing* loops. Engineers use negative feedback to stabilise systems in the presence of uncertainty.

System dynamics research has shown that in the case of socio-economic systems at least, many patterns of behaviour are generated by a small set of simplified 'generic structures'. Various classifications have been made [7]; see for example, generic infrastructures [8]. System archetypes are causal loop representations of generic patterns of behaviour over time, and are particularly useful for illustrating counter-intuitive behaviour [9]. Like all models, system archetypes are merely *approximations* of systems and their behaviour. Their value arises from the

compelling way in which they convey system insights [10]. We believe, in particular, that they provide important insights into accident causation in socio-technical systems.

Building Blocks: System dynamics models are built from three building blocks: the reinforcing loop, the balancing loop, and the delay.

A *Reinforcing Loop* is a structure that feeds on itself to produce growth or decline (positive feedback). An increase in *Variable 1* leads to an increase in *Variable 2*, as indicated by the '+' sign, which in turn leads to an increase in *Variable 1*, and so on. In the absence of external influences, both *Variable 1* and *Variable 2* will grow or decline exponentially. A characteristic of exponential growth is that the doubling time is constant. Because the initial growth is slow, it may be unnoticed until it becomes rapid, at which point it may be too late to control the growth. Reinforcing loops "generate growth, amplify deviations, and reinforce change" [6].

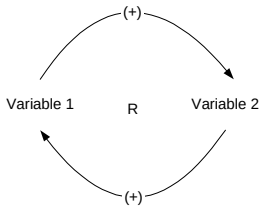


Figure 1 – Causal Loop Diagram of the Reinforcing Loop

A *Balancing Loop* is a structure that attempts to move a *variable* value to a *desired or reference value* through some *action* (negative feedback). The difference between the current state and the desired state is perceived as an *error*. An action proportional to the error is taken to decrease the error, so that, over time, the current state approaches the desired state. While the reinforcing loop tends to display exponential growth or decline, the balancing loop tends to settle down to the desired state. Because the size of the remedial action is proportional to the size of the error, the current state initially rapidly approaches the desired state. As the error decreases, the rate at which the current state approaches the desired state decreases.

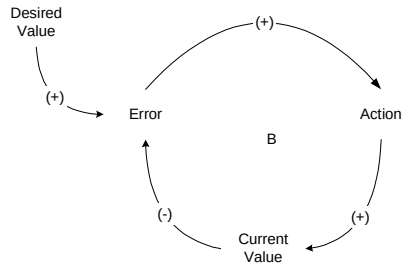


Figure 2 – Causal Loop Diagram of the Balancing Loop

Delays are used to model the time that elapses between cause and effect, and are indicated by a double line (Figure 3). Delays make it difficult to link cause and effect (dynamic complexity) and may result in unstable system behaviour. Consider the problem of navigating a ship down a narrow channel. Suppose that the ship is veering to one side of the channel, and the helmsman wishes to correct the course. Due to the ship's inertia, adjusting the rudder will not result in an immediate course change. There is a *delay* between a change in the rudder position and the resulting course change. In stressful situations, even experienced helmsmen may interpret a

delayed response as a complete lack of response, and accordingly make a larger change in the rudder position. When the ship's inertia is eventually overcome, the helmsman finds himself sailing towards the opposite side of the channel. If the helmsman continues to over-correct in this way, the ship will veer wildly from one side of the channel to the other, and may run aground.

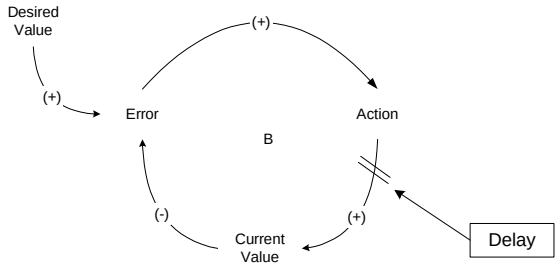


Figure 3 – Causal Loop Diagram of the Balancing Loop with Delay

Toward a Set of Safety Archetypes

In this section we propose a preliminary set of safety archetypes. General system behavioural archetypes have been described by Braun [11] and Wolstenholme [7]. While the general archetypes apply to all behaviour, our safety archetypes address specific behaviour related to flaws in an organization's safety culture. These safety archetypes can assist in representing and understanding the dynamic forces behind accidents and help accident investigators in their search for causal factors.

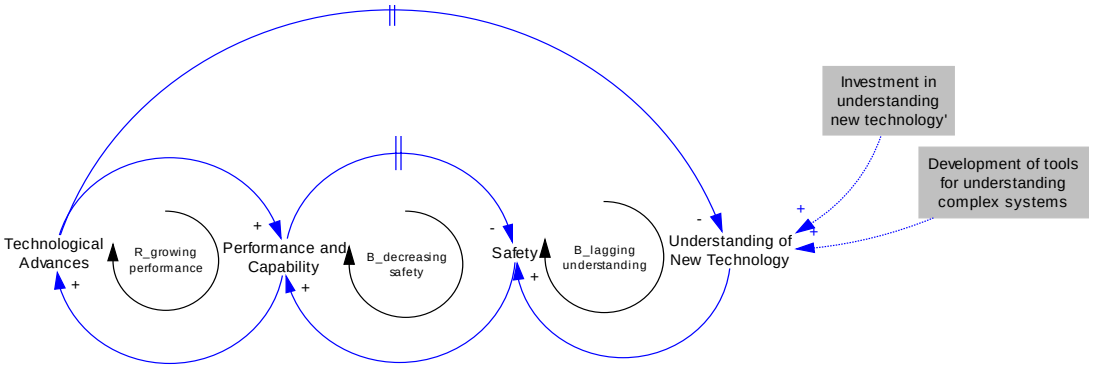


Figure 4 – Stagnant Safety Practices in the Face of Technological Advances

Stagnant Safety Practices in the Face of Technological Advances: This structure consists of a reinforcing loop ($R_{growing\ performance}$) and a balancing loop ($B_{decreasing\ safety}$). The reinforcing loop consists of some action in one part of an organisation, intended to achieve some outcome. Initially, the action is successful ($R_{growing\ performance}$), but after a time a constraint on performance is reached and the system reacts to limit the outcome ($B_{decreasing\ safety}$).

Here the constraint on safety is our understanding of new technology and the systems in which it is embedded. Technological advances result in an increase in performance in many areas, which in turn drives more advances ($R_{growing\ performance}$). As the speed of change accelerates, understanding of the safety implications lags further behind ($B_{lagging\ understanding}$). A characteristic feature of modern systems is that their complexity often exceeds our grasp. For example, it is alarmingly easy to write software whose behaviour cannot be predicted under all circumstances. This lack of

understanding translates into a decrease in safety ($B_{decreasing\ safety}$). We can ameliorate the problem by investing more resources in our understanding of new technologies, and by developing tools for understanding complex systems.

Decreasing Safety Consciousness: The success of a safety program may be limited by the characteristics of the system to which the program is applied, or by the nature of the program itself. A strategy, policy, or process that initially promotes improved safety may eventually reach a point where its continued application may cause a decline in safety.

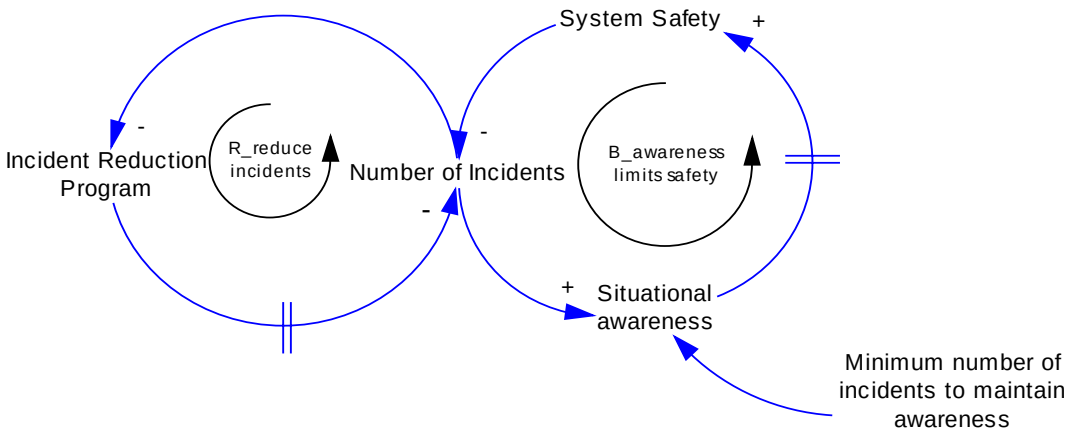


Figure 5 – *Decreasing Safety Consciousness*—Incident reduction measures initially improve system safety ($R_{reduce\ incidents}$). But the absence of incidents renders the system mute, and situational awareness of safety is decreased. The result is a decrease in system safety ($B_{awareness\ limits\ safety}$).

Consider the case of ultra-safe systems[1]. Common sense tells us that in order to increase safety, errors, incidents and breakdowns must be reduced or eliminated. This is true for systems where the rate of incidents and accidents is high. In the case of ultra-safe systems, continued elimination of errors, incidents, and breakdowns may paradoxically decrease safety [12].

Continued optimisation of a given set of safety measures, does not necessarily increase safety further. Over-optimisation numbs the adaptive capabilities of human and technical systems, while covering up minor system failures. In the case of error reduction, for example, it has been found that error plays an ecological role in the control of performance, and that detected errors are necessary to maintain situational awareness. Similarly, programs to reduce the number of incidents and breakdowns may also perversely decrease safety. As the perceived level of safety increases, investments are redirected from safety measures to improving system performance. Over-stretched system performance leads to new risks, which may materialise in the form of disastrous accidents [2]. Beyond a certain incident reduction quota, the absence of incidents, as opposed to the presence of a minimum number of incidents, does not prevent accidents from occurring. It may be necessary to tolerate a certain level of errors, incidents, breakdowns, and even accidents to protect the system against disastrous accidents. Figure 5 illustrates how incident reduction programs can result in a decrease in system safety.

[1] Amalberti defines ultra-safe systems as those where the risk of disaster is below one accident per 10^7 events [12]. For this discussion, it is sufficient to define high-risk systems as those that are not ultra-safe.

Amalberti argues that the combination of a system with a given set of safety measures bears within itself a maximum safety potential, which cannot be exceeded by continued optimisation of those safety measures. Continued optimisation of a particular safety measure 'mutes' some system aspects, thereby decreasing system awareness and adversely affecting safety. To obtain further increases in safety beyond this limit, additional, new safety measures are necessary. Therefore, to maintain safety, safety measures must be aggregated, but no single safety measure should be overly optimised.

Consider the strong emphasis on redundancy as a safety and reliability measure in many systems. Some degree of redundancy is useful in increasing reliability, and possibly safety. But more redundancy is not necessarily better, and may be worse. While redundancy may increase reliability, it does not necessarily increase and may decrease safety. First, a reliance on redundancy may lead to decreased emphasis on other safety engineering techniques. If system designers believe that redundancy will limit the effect of design errors they may be less motivated to find and eliminate these errors. In practice, redundancy may 'cover up', or mute, design errors and prevent them from becoming visible until something catastrophic occurs. Second, increasing redundancy increases system complexity. More complex systems are less amenable to testing and maintenance, and their properties and behaviour are difficult to predict accurately [13].

For example, an Air Force system included a relief valve to be opened by the operator to protect against over-pressurisation [4]. A secondary valve was installed as backup in case the primary relief valve failed. The operator had to know when the primary valve had not opened in order to determine that the secondary valve had to be opened. One day, the operator issued a command to open the primary valve. The position indicator and open indicator lights both illuminated although the primary relief valve had not opened. The operator, thinking that the primary valve had opened, did not activate the secondary valve and an explosion occurred. A post-accident investigation discovered that the indicator light circuit was wired to indicate only the presence of power at the valve, and not the actual valve position. The indicator showed only that the activation button had been pushed, not that the valve had opened. Redundancy could not provide protection against the underlying design error. Worse, the overconfidence provided by the redundancy convinced the engineers that an examination of the wiring design was not needed and the design error was therefore not found.

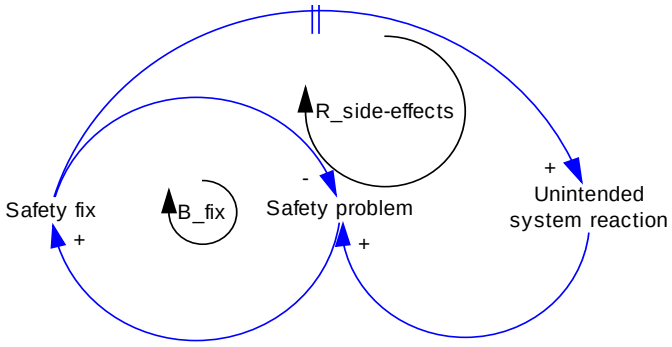


Figure 6 – Unintended Side Effects of Safety Fixes

Unintended Side Effects of Safety Fixes: The unintended consequences of poorly designed responses to safety problems, whether they are symptomatic treatments or supposed fundamental solutions, can worsen the problem.

This structure consists of a balancing loop (B_{fix}) and a reinforcing loop ($R_{side\text{-}effects}$). The loops interact so that the desired result initially produced by the safety fix in the balancing loop is, after some delay, offset by the undesired side effects in the reinforcing loop. Initially, the *Safety fix* ameliorates the *Safety problem* (B_{fix}). After a delay, the *Unintended system reaction* becomes visible ($R_{side\text{-}effects}$). Undesired aspects of the system reaction worsen the problem, and accordingly the safety fix is applied more strongly ($R_{side\text{-}effects}$). The safety fix ironically contributes to the worsening of the problem.

Well-intentioned, commonplace solutions to safety problems often fail to help, have unintended side effects, or exacerbate problems. The example below illustrates how disciplining workers and writing more detailed procedures may fail to reduce the number of equipment breakdowns.

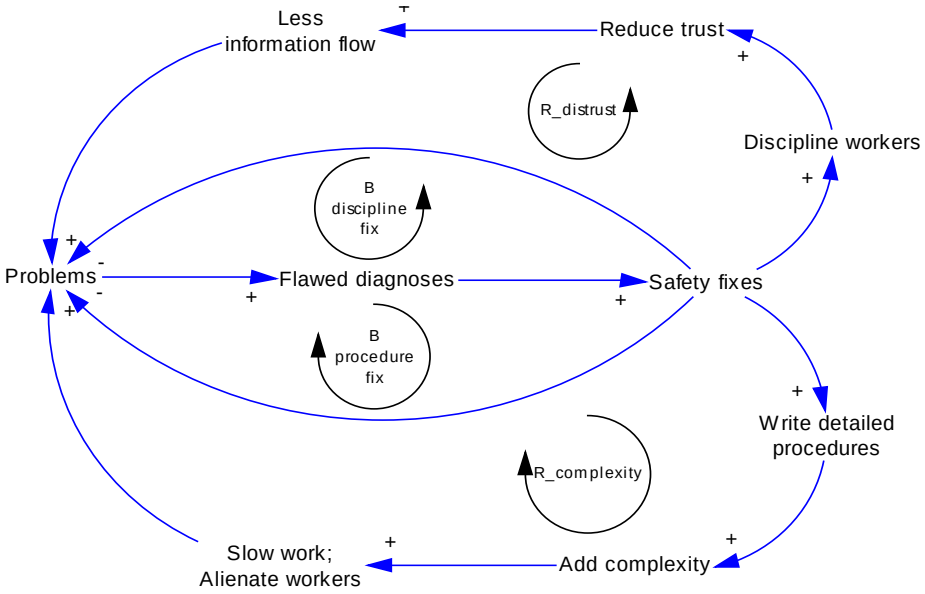


Figure 7 – The common response to incidents or accidents of disciplining workers or writing more detailed procedures is intended to solve the problems (balancing loops $B_{discipline\ fix}$ and $B_{procedure\ fix}$). But these fixes often result in reinforcing loops ($R_{distrust}$ and $R_{complexity}$) that eventually make the problems worse.

Consider a plant that is experiencing increasing equipment breakdowns, which are attributed to poor maintenance. A typical 'fix' for maintenance-related problems is to write more detailed maintenance procedures and to monitor compliance with these procedures more closely. More detailed procedures can translate to fewer errors in a particular task. But workers also tend to view more detailed procedures and closer supervision as mistrust and regimentation, causing them to lose motivation, or comply blindly or maliciously with procedures that may be incomplete or incorrect. Skilled workers may find the new regime intrusive and look for more interesting work elsewhere. Excessive restrictions on behaviour discourages problem solving and encourages blind adherence to procedures, even when such compliance is not optimal in terms of safety or productivity. Blaming or disciplining individual workers, designed to create an atmosphere of accountability, encourages all workers to hide problems. For example, when the Federal Aviation Administration provided immunity from prosecution to pilots who reported near-collisions, the number of reports tripled; when immunity was later retracted, the number of reports decreased six-fold [14]. When incidents are deliberately concealed, the underlying problems do not become visible, often worsen, and may lead to more problems (Figure 7).

Unintended Side Effects behaviour occurs when the fundamental problem is not understood, or when the solutions to the fundamental problem are not appropriate or are improperly implemented. We can avoid or escape this behaviour by correctly identifying the fundamental problem and designing appropriate solution strategies. Identifying the fundamental problem is often difficult, and designing and implementing solution strategies can be challenging. An awareness of the long-term negative implications that fixes often have can provide the impetus to search for fundamental solutions instead.

Fixing Symptoms Rather Than Root Causes: In this archetype, *Symptomatic solutions* are implemented in response to *Problem Symptoms* ($B_{symptoms}$), temporarily decreasing the symptoms (Figure 8). If the *Fundamental Solution* is known, *Side Effects* of the symptomatic solutions may either decrease the desire to implement the fundamental solution, or act to decrease the effectiveness of the fundamental solution ($R_{side\ effects}$). Alternatively, if the fundamental solution is not known, the symptomatic solutions may decrease the ability to find the fundamental solution, for example by masking the problem symptoms.

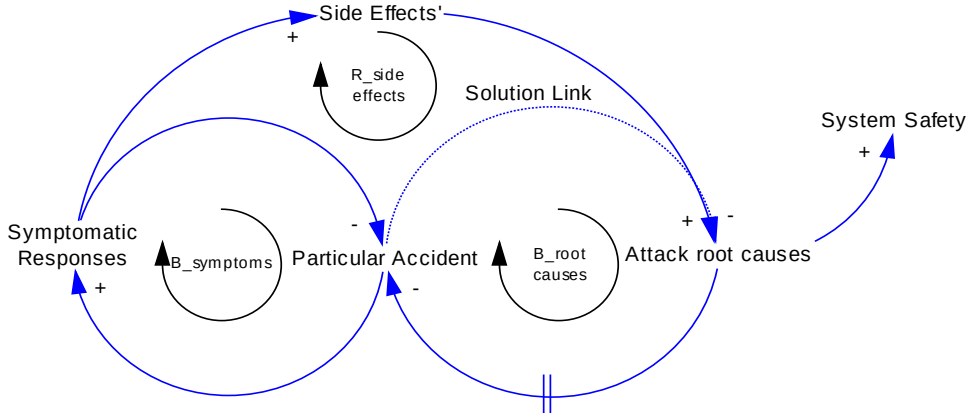


Figure 8 – Fixing Symptoms: Symptomatic solutions decrease the likelihood of recurrence of the same accident ($B_{symptoms}$), but do not address the underlying conditions that allowed the accident to occur in the first place. A side effect is that the impetus to find fundamental solutions is decreased ($R_{side\ effects}$). Organisations should instead perform root cause analysis and use the resulting insights to formulate fundamental solutions that address the underlying systemic causal factors ($B_{root\ causes}$).

Fixing Symptoms illustrates the tension between the appeal of short-term, symptomatic solutions, and the long-term impact of fundamental solutions. Symptomatic solutions are usually easier, faster, and cheaper to implement than long-term fundamental solutions. Initially, positive results to symptomatic solutions are seen immediately, as the visible symptoms are eliminated. Once a symptomatic solution has been successfully applied, the pressure to find and implement a fundamental solution tends to decrease. Over time, the solutions may become less effective, or different symptoms of the underlying problem may arise; in response new symptomatic solutions are devised. The underlying problem remains. If the fundamental problem is not dealt with, symptoms can be expected to continue surfacing in various forms. Long-term, fundamental solutions, on the other hand, may be more difficult to devise, more difficult to implement, take longer to show results, and are often initially more costly. At the same time, external pressures often demand a 'quick-fix' to the problem.

The reactive focus of many safety programs results in placing primary emphasis on investigating previous incidents and accidents in an attempt to prevent future accidents. These efforts are not

always fruitful. Excessive focus is placed on preventing recurrence of exactly the same accident, without taking sufficient account of the underlying systemic factors that allowed safety to deteriorate [4]. Attempts to identify the deeper factors or conditions that allowed the accident to occur (i.e., root cause analysis) are often insufficient.

For example, Carroll et al. have identified instances of inadequate root cause analysis at nuclear plants [15]. In the nuclear and chemical industries, problem investigation teams are assigned to examine serious incidents and troubling trends. These investigations are part of corrective action programs to improve safety and performance. Although considerable resources are devoted to these programs, the investigations do not always result in effective learning. The investigations studied by the authors tended to focus on only a few proximal causes. These causes were typically technical or involved human error, and their solutions were obvious, easily implemented, and acceptable to powerful stakeholders. Little effort was made to uncover root causes or devise fundamental solutions.

Symptomatic solutions to accidents often only decrease the likelihood of that particular accident recurring. They do not eliminate the deeper structural deficiencies that led to the accident in the first place. Once a symptomatic solution has been successfully applied, the perceived need to solve the underlying structural problem may disappear, reducing the pressure to find a fundamental solution. To improve safety in the long term the fundamental problem or structural deficiency that is causing the symptoms must be identified.

For example, if an aircraft rudder failure is shown to be the result of insufficient or poor maintenance, the recommended action may be to improve the rudder maintenance procedures. But deeper problems, such as subtle management pressure to increase maintenance throughput, may have caused the maintenance to be poorly performed in the first place.

Identifying the root causes of incidents and accidents is not always easy to do. Symptomatic solutions may be suppressing the symptoms, creating the illusion that no problem exists. These solutions may be consciously or unconsciously formulated and applied. Unconsciously applied solutions (e.g. unconsciously correcting for misaligned steering on a motor vehicle) may so successfully mask the underlying problem that operators are not aware of the problem symptoms, let alone the fundamental problem. In order to understand the symptoms of the problem, it is necessary to identify the conscious and unconscious symptomatic solutions. Because any individual only has a limited view of the system, obtaining different viewpoints of the symptoms, the problem, and the system can help in identifying the fundamental problem.

Eliminating root causes is likely to be more difficult, time-consuming, and costly to implement than implementing symptomatic solutions. It is essential to obtain commitment from all parties involved with the implementation of the proposed solution. Without such commitment, the solution is unlikely to be successfully applied. Side effects of the solution must be identified as far as possible. Of course it may not be possible to foresee all the side effects. Awareness of the potential for side effects makes it easier to identify and deal with them if they do occur. Where side effects of symptomatic solutions may undermine the fundamental solution, it is necessary to stop applying these solutions before applying the fundamental solution.

Eroding Safety: This archetype illustrates how safety goals may erode or become subverted over time. We can expect to observe *Eroding Safety* behaviour in systems where an accident was preceded by a declining emphasis on safety, such as decreasing safety goals. This decline is an example of migration towards unsafe behaviour. *Eroding Safety* is often difficult to observe while it is occurring because change tends to happen gradually. At short time scales, changes may be

imperceptible. It is only after an accident has occurred that the extent of change is noticed, if at all. The first example illustrates how complacency can grow in an organisation with a history of safe operation. The second example illustrates why well-designed safety programs do not always achieve their goals.

Complacency: A history of safe operations often results in growing complacency. Figure 9 illustrates how complacency can arise. Consider a system that initially operates with a high accident rate. In order to bring the accident rate down, the system is closely monitored, possibly both internally (company rules and procedures) and externally (government regulation). Close oversight eventually decreases the accident rate, and may bring it to the point where people do not believe that accidents can or will occur. In the apparent absence of a threat to safety, oversight may seem draconian and unnecessarily costly. Coupled with budgetary pressures, this anti-regulation sentiment creates pressure to decrease oversight. Decreased oversight is manifested on the one hand by less training and fewer or less strict certification requirements, and on the other hand by decreased inspection and monitoring. A decrease in these activities eventually leads to an increase in the risk of accidents, and so the accident rate increases.

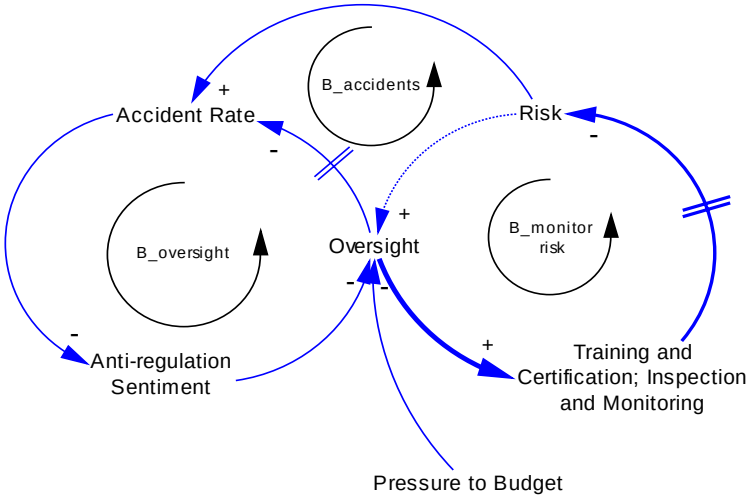


Figure 9 – Complacency occurs when low accident rates encourage anti-regulation sentiment, which coupled with budget pressures leads to less oversight ($B_{oversight}$). But decreasing oversight means decreased training and certification, and decreased inspection and monitoring, which in turn increases the risk of accidents and hence the accident rate ($B_{accidents}$). One way to avoid the complacency trap is to continuously monitor risk and set the level of oversight accordingly ($B_{monitor\ risk}$).

Following the Apollo launch pad fire in 1967, NASA established one of the best system safety programs of the time [16]. But nearly two decades later the Rogers Commission report on the Challenger accident referred to a 'Silent Safety Program' that had lost some of its effectiveness since Apollo. In particular, the report cited growing complacency at the agency, as the perception grew that Shuttle operations were routine (emphasis added) [17]:

> Following successful completion of the orbital flight test phase of the Shuttle program, the system was declared to be operational. Subsequently, several safety, reliability and quality assurance organizations found themselves with reduced and/or reorganized functional capability… The apparent reason for such actions was *a perception that less safety, reliability and quality assurance activity would be required during 'routine' Shuttle operations*. This reasoning was faulty. The machinery is highly complex, and the requirements are exacting… As the system

matures and the experience changes, careful tracking will be required to prevent premature failures… *Complacency and failures in supervision and reporting seriously aggravate these risks.*

The problem with complacency is twofold. First, it is difficult not to become complacent when success follows upon success. Second, it is difficult for an organisation to realise that it is becoming complacent, and often a serious accident is required to shake the complacency.

Organisations can avoid sinking into complacency by continuously monitoring risk and setting the level of oversight accordingly, as shown by the dotted line in Figure 9. Complacency arises because the accident rate usually does not immediately increase following a decrease in oversight. Inertia in the system temporarily keeps the accident risk at a low level, creating the impression that oversight is set at the appropriate level. All the while, the system is migrating towards the boundary of safe behaviour [2]. When accidents start occurring, the link to decreased oversight is not immediately obvious. When making the connection between risk and the level of oversight, the long-term trend in the risk level must be considered, rather than short-term fluctuations.

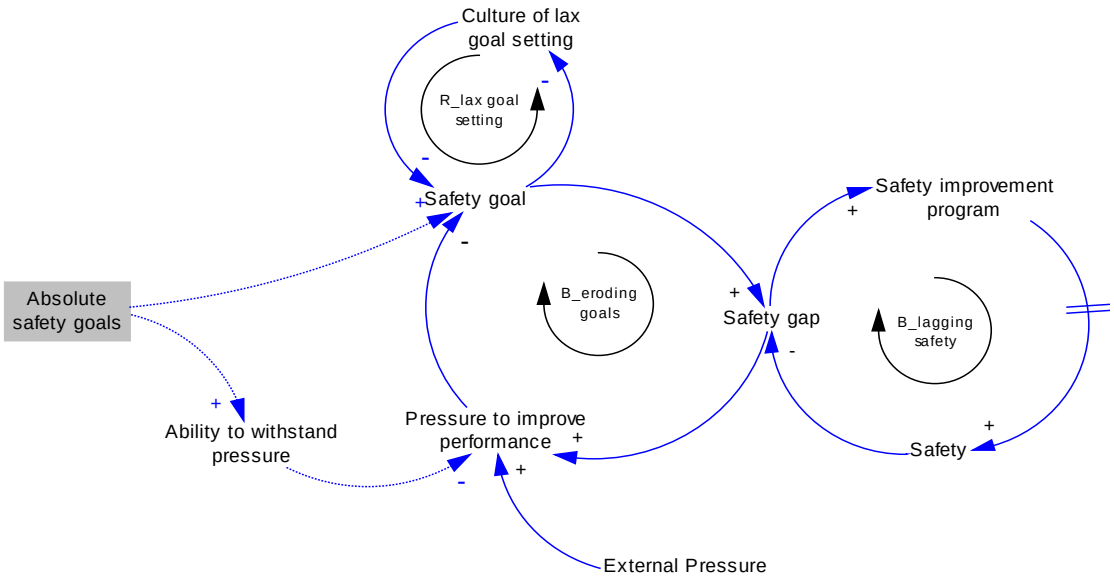


Figure 10 – Safety improvement programs typically do not show immediate results ($B_{lagging\ safety}$). External pressure results in a decreasing emphasis on safety and a lowering of the safety goals ($B_{safety\ emphasis}$ and $B_{eroding\ goals}$). These balancing loops interact to repeatedly lower the safety goal. Repeated lowering of the goal results in a reinforcing dynamic ($R_{lax\ goal\ setting}$) that encourages lax goal setting in the future. The problem can be addressed by setting absolute safety goals, perhaps based on some external standard.

Disappointing Safety Programs: *Eroding Safety* illustrates why safety programs do not always live up to their expectations (Figure 10). Safety improvement programs can be expensive and often do not show immediate results. While the eventual costs of not improving safety can be high, the immediate cost of a safety program is subject to external pressures (e.g. management pressure for performance improvement). The combination of seeming ineffectiveness and external pressures makes it tempting to adjust the goals of the safety program. This adjustment is not necessarily seen as a failure, and may even be viewed as an improvement.

For example, a common response to failed programs is to restructure parts of, or the entire organisation in question. After the Challenger accident NASA responded by reorganising the

safety and quality programs at NASA Headquarters and the field centres. A new office of Safety, Reliability, Maintainability and Quality Assurance (SRM&QA) was established and overall management of the safety function was elevated to the level of associate administrator, in an attempt to increase awareness of significant safety and quality issues at the highest levels of NASA management. This reorganisation was presented as one of the most significant improvements following the Challenger accident [18]. In fact, this reorganisation failed to achieve its goal over the long term, and many of the same 'silent safety program' characteristics have become evident following the Space Shuttle Columbia accident.

While restructuring and reorganisation is sometimes necessary, it does not always address the underlying problem. Another, more subtle form of downward goal adjustment is the eternally receding deadline. In this case, the goals remain the same, but the deadline for meeting the goals is continually shifted back, effectively lowering the goals.

Pressure for increased performance (e.g. delivery times, profit) can make it difficult to remain focussed on safety goals. *Eroding Safety* illustrates how these pressures can contribute to safety improvement goals not being met. The challenge is to resist external pressures that work against safety improvement programs, whether overtly or in a less obvious manner. Anchoring the safety goals to externally generated and enforced standards or deadlines can make adjustments in goals more visible or more difficult to make. For example, government regulators impose certain minimum safety standards on some industries, such as the nuclear power industry.

The safety program must provide a clear plan and a realistic timeframe for improving safety. It must provide concrete steps towards achieving the safety goal, as well as interim measures of progress. If a safety program is seen as working against performance (e.g. preventing on-time delivery of goods), there will be a reciprocal tendency to work against the program, thereby decreasing its effectiveness. Managers who pay lip service to safety programs but simultaneously demand increased performance encourage a lax attitude to safety at lower organisational levels. Only when there is buy-in at all levels of the organisation can a safety program succeed.

Incident Reporting Schemes: Consider what often happens when incident reporting schemes are implemented (Figure 11). The primary purpose of these schemes is to encourage workers to be more careful on a day-to-day basis, thus reducing the number of incidents. As an incentive to reduce the number of incidents, workers with the best safety records (as measured by fewest reported incidents) are rewarded. Rewarding workers who report the fewest number of incidents is an incentive to withhold information about small accidents and near misses. Underreporting of incidents creates the illusion that the system is becoming safer, when, in fact, it has merely been muted. Management becomes less aware of the behaviour of the system, and safety may therefore decrease. At the worker level, the original goal of increasing safety is subverted into one of reporting the fewest incidents. Ironically, the introduction of an incident reporting scheme can decrease safety, as found in a study of the California construction industry [19].

The *Eroding Safety* archetype illustrates how unforeseen side effects of safety programs can work against the success of the programs. In implementing safety programs it is essential to consider carefully what incentives or rewards will be used to ensure compliance. If symptomatic behaviour is rewarded (e.g. fewest reported incidents), it is likely that workers will find other ways to generate the same symptoms (e.g. underreporting incidents). If incentives are inappropriately formulated, compliance with the intent of the program may be lower than if no incentives were offered. This behaviour can also be observed in organisations that operate according to process certification standards. In this case the purported rewards are often not visible and employees view the requirements as impeding their normal working processes. Employees therefore obey

the letter of the process and documentation standards, but do not comply with the underlying intentions.

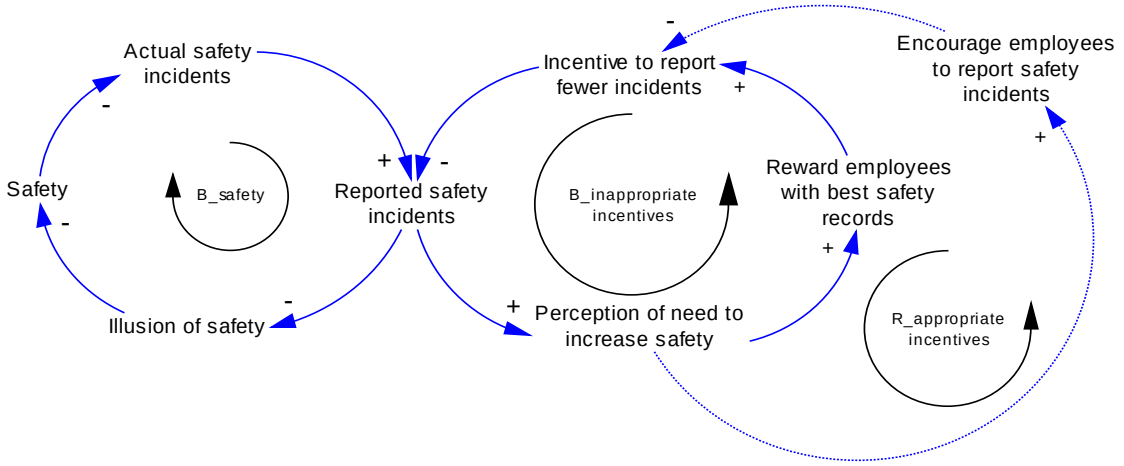


Figure 11 – An incident-reporting scheme is implemented to encourage workers to be more careful on a day-to-day basis. Workers with the best safety records are rewarded. The reward acts as an incentive to underreport incidents ($B_{inappropriate\ incentives}$). Underreporting of incidents decreases awareness of the system and creates the illusion of safety. The result is an unnoticed decrease in the system's safety (B_{safety}). One way of avoiding this type of behaviour is to encourage employees to report safety incidents, rather than rewarding employees with the best safety records ($R_{appropriate\ incentives}$).

The intent of safety programs must be communicated at all levels of the organisation. Employees must be provided with the necessary resources to perform their part in the programs. They must be empowered to make safety-based choices in cases where such decisions might adversely affect productivity. If employees understand the intent of, and are therefore committed to the program, they are more likely to comply with the intent than with the letter of the law.

Conclusions and Future Work

We have proposed a preliminary set of safety archetypes by specializing general system archetypes developed in system dynamics. The archetypes can be used to describe flaws in an organization's safety culture. In accident analysis, the archetypes can be used to model the dynamic aspects of safety-related behaviour at the organisational level. They are also useful in structuring post-investigation recommendations, by highlighting the mechanisms or root causes that led to the accident. The archetypes explain why safety-related decisions do not always result in the intended outcomes, and how independent decisions in different parts of an organisation can inadvertently interact to decrease safety.

In future work we will further develop this preliminary set of archetypes and demonstrate the application of the safety archetypes to accident analysis, by using them in the modelling of socio-technical accidents. In related work, they will be used in the creation of new approaches to risk assessment and management.

Acknowledgement

This research was supported in part by NSR ITR Grant CCR-0085829 and NASA Ames (Engineering for Complex Systems) Grant NAG2-1543.

References

[1] Perrow, Charles, Normal Accidents: Living with High-Risk Technologies, Princeton University Press, New Jersey, 1999.

[2] Rasmussen, Jens, "Risk Management in a Dynamic Society: A Modelling Problem," Safety Science, Vol. 27, No. 2, pp. 183-213, 1997.

[3] Leveson, Nancy G., Allen, Polly, Storey, Margaret-Anne, "The Analysis of a Friendly Fire Accident using a Systems Model of Accidents," *Proceedings of the 20th International System Safety Conference*, Denver Colorado, 5-9 August 2002.

[4] Leveson, Nancy G., *A New Approach to System Safety Engineering*, 2003. Available online at http://sunnyday.mit.edu.

[5] Leveson Nancy G., Daouk, Mirna, Dulac, Nicolas, and Marais, Karen, "Applying STAMP in Accident Analysis," Submitted to the *2nd Workshop on the Investigation and Reporting of Accidents*, September 2003.

[6] Sterman, John D., "System Dynamics: Systems Thinking and Modelling for a Complex World," *Proceedings of the ESD Internal Symposium*, MIT, Cambridge, MA, May 2002.

[7] Wolstenholme, Eric F., "Toward the Definition and Use of a Core Set of Archetypal Structures in System Dynamics," *System Dynamics Review*, Vol. 19, No. 1, Spring 2003, pp. 7-26.

[8] Paich, M., "Generic Structures," System Dynamics Review, Vol. 1, pp. 126-132, 1985.

[9] Senge, P. M., *The Fifth Discipline: The Art and Practice of the Learning Organisation*, Doubleday Currency, New York, 1990.

[10] Lane, David C., "Reinterpreting 'Generic Structure': Evolution, Application and Limitations of a Concept," *System Dynamics Review*, Vol. 12, pp. 87-120, 1996.

[11] Braun, William, *The System Archetypes*, 2002. Available online at: http://www.uni-klu.ac.at/~gossimit/pap/sd/wb_sysarch.pdf.

[12] Amalberti, R., "The Paradoxes of Almost Totally Safe Transportation Systems," *Safety Science*, Vol. 37, pp. 109-126, 2001.

[13] Graham, John, *Fast Reactor Safety*, Academic Press, New York, 1971.

[14] Tamuz, M., "Developing Organizational Safety Information Systems." In Apostolakis, George E., and Wu J.S. (Eds.), *Proceedings of PSAM II, Vol. 2, Los Angeles*, University of California, pp. 71: 7-12.

[15] Carroll, John S., Rudolph, Jenny W. and Hatakenaka, Sachi, "Organizational Learning from Experience in High-Hazard Industries: Problem Investigations as Off-line Reflective Practice", MIT Sloan Working Paper No. 4359-02, April 2002. Available online at: http://ssrn.com/abstract=305718

[16] Commission on Engineering and Technical Systems, *An Assessment of Space Shuttle Flight Software Development Processes*, Committee for Review of Oversight Mechanisms for Space Shuttle Flight Software Processes, Aeronautics and Space Engineering Board, National Research Council, National Academies Press, 1993. Available online at: http://www.nap.edu/

[17] Rogers, William P., Chairman, *Report of the Presidential Commission on the Space Shuttle Challenger Accident*, Government Printing Office, Washington DC, 1986.

[18] Press Release 88-01-05. Available online at: http://spacelink.nasa.gov/NASA.News/NASA.News.Releases/.index.html

[19] Levitt, Raymond E., and Henry W. Parker, "Reducing Construction Accidents - Top Management's Role," *ASCE Journal of the Construction Division*, Vol. 102, No. CO3, September 1976, pp. 465-478.

Biography

Karen Marais, MIT Department of Aeronautics and Astronautics; Cambridge, Massachusetts, U.S.A.; telephone +1.617.258.5046; e-mail karen.marais@alum.mit.edu.

Ms. Marais is a doctoral candidate at MIT in the Department of Aeronautics and Astronautics. Her research interests include systems engineering, safety, and risk assessment.

Nancy Leveson, MIT Department of Aeronautics and Astronautics; Cambridge, Massachusetts, U.S.A., telephone +1.617.258.0505; email leveson@mit.edu.

Dr. Leveson has dual appointments as a professor of Aeronautics and Astronautics and a professor of Engineering Systems at MIT. She is a member of the National Academy of Engineering. Her research interests include system engineering, system safety, human-computer interaction, and software engineering.

Probabilistic Causal Analysis for System Safety Risk Assessments in Commercial Air Transport

James T. Luxhøj, Ph.D.; Department of Industrial and Systems Engineering, Rutgers University;
96 Frelinghuysen Road, Piscataway, NJ 08854-8018 USA

Keywords: Aviation Safety, Risk Management

Abstract

Aviation is one of the critical modes of our national transportation system. As such, it is essential that new technologies be continually developed to ensure that a safe mode of transportation becomes even safer in the future. The NASA Aviation Safety Program (AvSP) is managing the development of new technologies and interventions aimed at reducing the fatal aviation accident rate by a factor of 5 by year 2007 and by a factor of 10 by year 2022. A portfolio assessment is currently being conducted to determine the projected impact that the new technologies and/or interventions may have on reducing aviation safety system risk. This paper reports on advanced risk analytics that combine the use of a human error taxonomy, probabilistic Bayesian Belief Networks, and case-based scenarios to assess a relative risk intensity metric. A sample case is used for illustrative purposes.

Introduction

Commercial air transportation in the United States is a complex array of many diverse, yet interrelated system components. There is a plethora of varied human, technical, environmental, and organizational factors that affect the performance of the National Airspace System (NAS). Through the years, numerous qualitative and quantitative approaches to aviation risk identification, modeling, and evaluation have been developed and have contributed in a seminal way to the understanding of aviation safety risk [1]. However, while methods exist for identifying aviation risk factors, there is a paucity of analytical methods for analyzing and interpreting the complex interactions of the various system risk factors. There has been a persistent need to develop advanced risk analytics that move beyond the essential identification of risk factors to enhanced system modeling and evaluation of complex causality as well as to assessing various combinations of risk mitigation strategies [2-8].

The NASA Aviation Safety Program (AvSP) Office located at the NASA Langley Research Center is managing the joint industry/government/university development of 48 new technologies/interventions intended to improve aviation system safety [9]. Figure 1 displays the principal categories of the new technologies. As an example, there are four NASA AvSP technologies focused at aircraft Loss of Control (LOC) issues: (1) Auto-configurable aircraft controls given upset or specific system failures (2) Database of upset/control and recovery phenomena that provides data for simulation training and/or decision-aids (3) Maintenance visualization support/training such that system failures are reduced (these failures could lead to LOC in the causal chain) and (4) Weather (Wx) decision-aids that evaluate and warn pilot of conditions (icing, turbulence, etc.) that may cause upset. There is a requirement to develop an analytical method that facilitates assessing the projected impact of the various technologies and/or interventions upon system risk reduction [10].

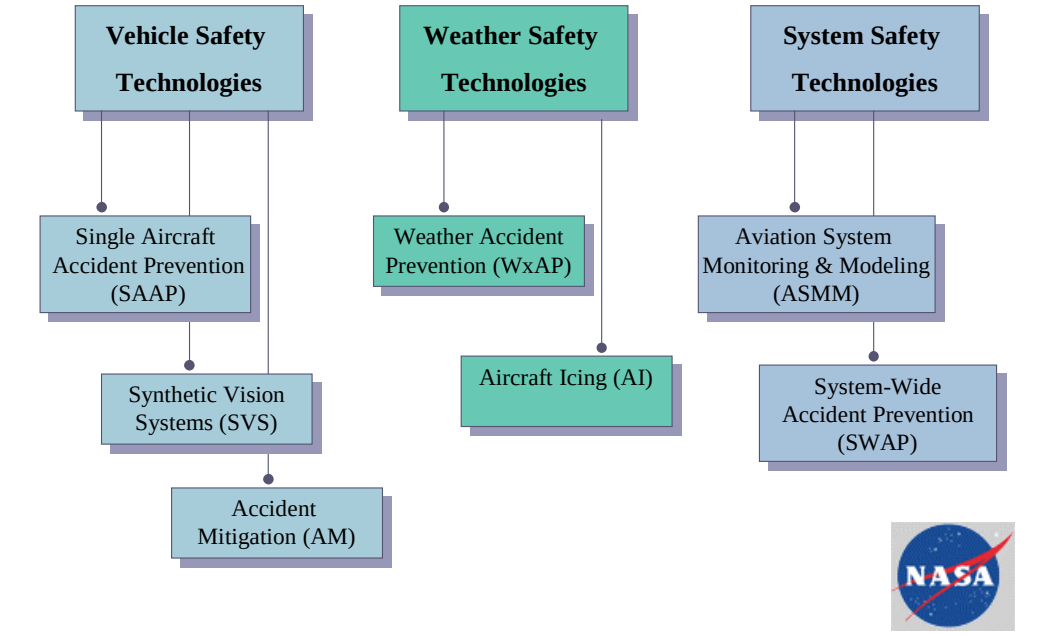


Figure 1 - Categorical Overview of NASA Technologies

Luxhøj, et al. [11-14] report on the development of an *Aviation System Risk Model (ASRM)* that uses the underlying probabilistic methodology of Bayesian Belief Networks (BBNs) and influence diagrams to graphically portray causal factor interactions. The ASRM uses the Information Technology (IT) embedded in the HUGIN BBN software as an enabling technology [14]. HUGIN is software to facilitate BBN modeling [15-17]. The ASRM is being modified to graphically portray a risk metric, termed the *relative risk intensity*, to illustrate perturbations from a baseline period. In Phase 1 of this research, aircraft accident scenarios, such as Loss of Control (LOC) and Maintenance (MAIN)-related, have been developed using the combined approach of *analytic generalization* from case studies [18] and from knowledge engineering sessions with subject matter experts (SMEs). van Vurren [19] uses a similar approach in his study of mishaps in the steel industry and medical domain.

Typically, 60%-80% of all accidents are attributed to human error. While mechanical, environmental, and operational factors are necessarily to be included in a full system precursor analysis, certainly, understanding human error is essential to aviation system risk modeling. In a recent research modification, the ASRM has been adapted to include the Human Factors Analysis and Classification System (HFACS) [20-22] taxonomy to expand the examination of human causal factors and to support the identification of safety risk intervention strategies. HFACS [21] focuses on human error modeling and includes organizational influences (resource management, organizational climate, and organizational processes), preconditions for unsafe acts (adverse mental states, adverse physiological states, physical/mental limitations, crew resources management, and personal readiness), and individual unsafe acts (decision errors, skill-based errors, perceptual errors, routine violations, and exceptional violations). While other general domain taxonomies exist for the classification of human error, such as the Eindhoven Classification Method (ECM) [19], HFACS is becoming widely disseminated in both military

and commercial organizations as a tool for understanding the role of human error in aviation accident analysis. Shappell and Wiegmann have led an effort to code all military, commercial and general aviation accidents in the United States using the HFACS framework. The database consists of over 16,000 aviation accidents involving human error and includes all the Parts 121 and 135, scheduled and non-scheduled, airline accidents since 1990. These data are being used in the ASRM to initially seed the model and to facilitate exercising the model.

The ASRM prototype methodology and tool have been through an initial testing and evaluation period and offer promise. However, additional analytical research and tool development are required in order to realize the full potential of this new type of system risk model. The joint research between the Federal Aviation Administration (FAA), NASA, and Rutgers University involves extending the ASRM to modeling the complex interactions of the many diverse human, technical, environmental, and organizational factors that affect commercial air transportation. The intent of the new research is to develop more comprehensive ASRM case studies reflective of multiple accident scenarios selected by the NASA Aviation Safety Program (AvSP). The case studies model the change in system safety risk within the accident scenarios due to risk mitigation strategies proposed by NASA AvSP by examining the impact that technology insertions and/or interventions may have on reducing the relative system safety risk. For example, it is envisioned that the size of the HFACS database will exercise the ASRM and yield a model capable of evaluating the NASA AvSP intervention/mitigation strategies focused on human factors before they are fielded. Eventually, other data sources dealing with mechanical, software, environmental, and operational risk factors will be integrated into the ASRM as these data sources become available.

Risk Analytics

Risk is a mathematical expression that has two components - *likelihood* and *severity*. Risk is an expression that attempts to answer two questions at the same time; how likely? and with what consequence? These components of risk can be defined as follows:

Hazard
A hazard implies any event that has the potential to produce an adverse outcome with respect to the system. The system can refer to a piece of equipment or a single engine or an entire airplane [23]. In most situations, the fact that a hazard is present does not necessarily mean that there will be an adverse outcome with respect to a system. A “successful” hazard is an event when a hazard attacks the system and results in damages.

Likelihood
Likelihood or probability represents the chance that a given hazard will lead to an adverse impact on a system. It can be described qualitatively as well as quantitatively.

Severity
Severity represents the damages that result in the case of successful hazard event. The severity could be measured in terms of dollar value, human life or etc. Risk, in its quantitative form, risk can be expressed as follows:

$$R = P * S$$

where R is risk, P is probability or likelihood of an event and S is severity.

Since it is not always possible to calculate the risk quantitatively or since it is sometimes necessary to assign qualitative descriptions to likelihood and damages, risk is frequently expressed in relative or qualitative terms. Figure 2 illustrates qualitative descriptions of likelihood, severity and risk developed by the FAA's Office of System Safety (ASY-300).

FAA Order 8040.4 establishes the safety risk management policy and prescribes procedures for implementing safety risk management as a decisionmaking tool within the FAA (see http://www.asy.faa.gov). This document provides general guidelines and principles for safety risk assessment and risk characterization.

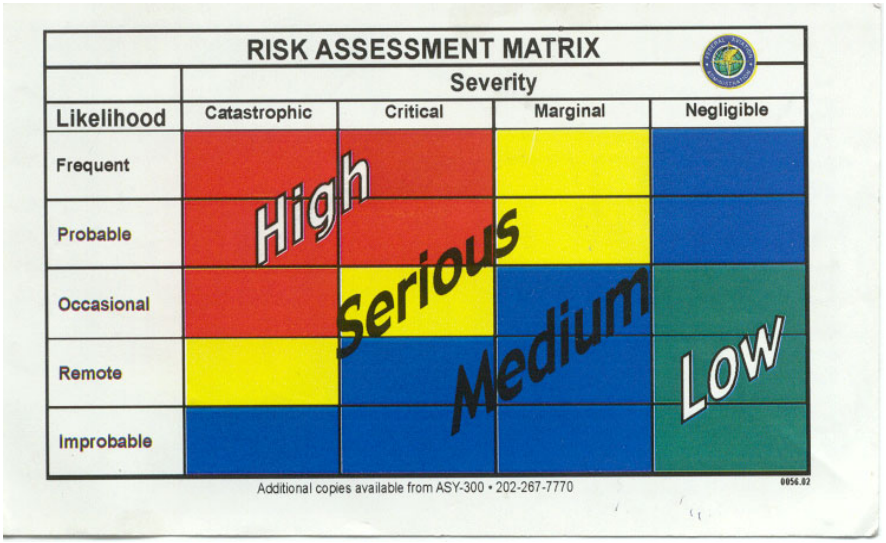


Figure 2 - Risk Matrix
(Source: ASY–300, Office of System Safety, FAA)

Relative Risk Intensity Metric: This research introduces the notion of a *relative risk intensity* metric. This metric uses the matrix cells of Figure 2 to create a visual profile of risk relative to some baseline. Note that the risk matrix in Figure 2 is a 5 x 4 matrix. Thus, from observing Figure 2, the following risk intensity levels are computed:

# of cells	Risk Intensity Level	
3/20	0%-15%	Low
8/20	15% - 55%	Medium
4/20	55% - 75%	Serious
5/20	75% - 100%	High

Figure 3 displays the use of the relative risk intensity combined with an influence diagram. The ASRM itself facilitates the computation of the likelihood portion of risk, or notionally, the "intensity". To fully understand risk, the ASRM software prototype is also being updated to enable the user to view a severity histogram per accident type (e.g., LOC, Controlled Flight Into Terrain (CFIT), etc.) as determined by data analysis.

Research Methodology

The underlying research methodology is comprised of three analytical approaches:

- the Human Factors Analysis and Classification System (HFACS)
- Bayesian Belief Networks (BBNs)
- case studies

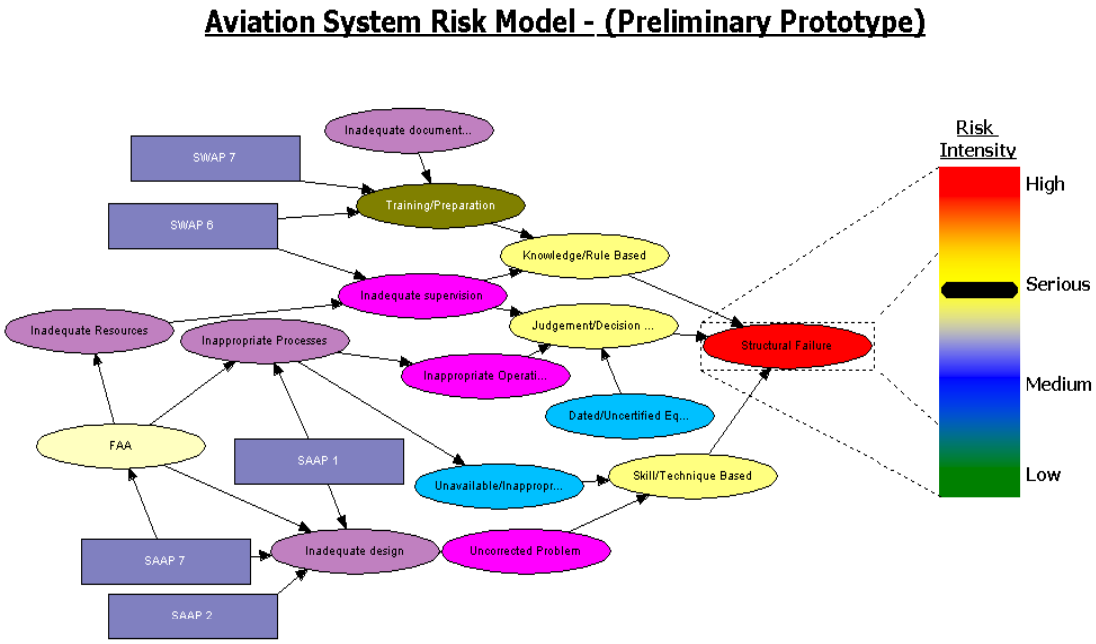


Figure 3 - Relative Risk Intensity Metric

The Human Factors Analysis and Classification System (HFACS): The HFACS is an analytical approach for classifying human error that is based on the Reason framework of system safety theory. While there are numerous contributing factors to aircraft accidents, such as operational, weather, etc., nevertheless, 60%-80% of all accidents are attributed to human error. The HFACS provides a fundamental analytical method for approaching causal modeling and the factors are illustrated in Figure 4. For a detailed description of the HFACS taxonomy, see [20].

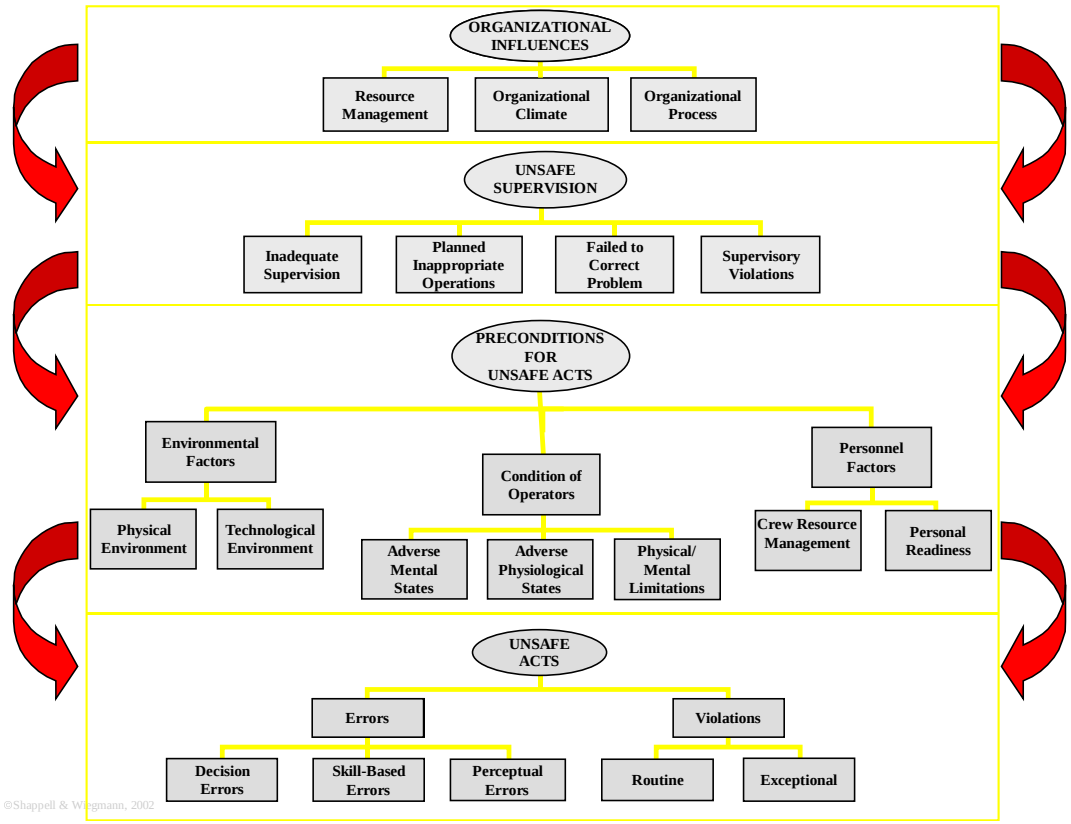


Figure 4 - The Human Factors Analysis and Classification System (HFACS) (Source: [20])

Bayesian Belief Networks (BBNs): One of the most important factors that should be taken into account when building models of accident causation is uncertainty. Probability theory derives solutions to the problem of reasoning under uncertainty in the face of limited information. In recent years, Belief Networks have been used as the main methodology for numerous tasks that involve reasoning under uncertainty. Belief networks provide efficient symbolic representations of probability models, together with the efficient inference algorithms for probabilistic reasoning [24-25].

Figure 5 displays an influence diagram with chance modes or variables represented as circles and decision nodes (see D1, D2, and D3) shown as rectangles. A decision variable can be related to one or multiple chance variables or multiple decision variables can be related to one particular chance variable. In this research, the decision nodes represent the AvSP technology and/or intervention insertions.

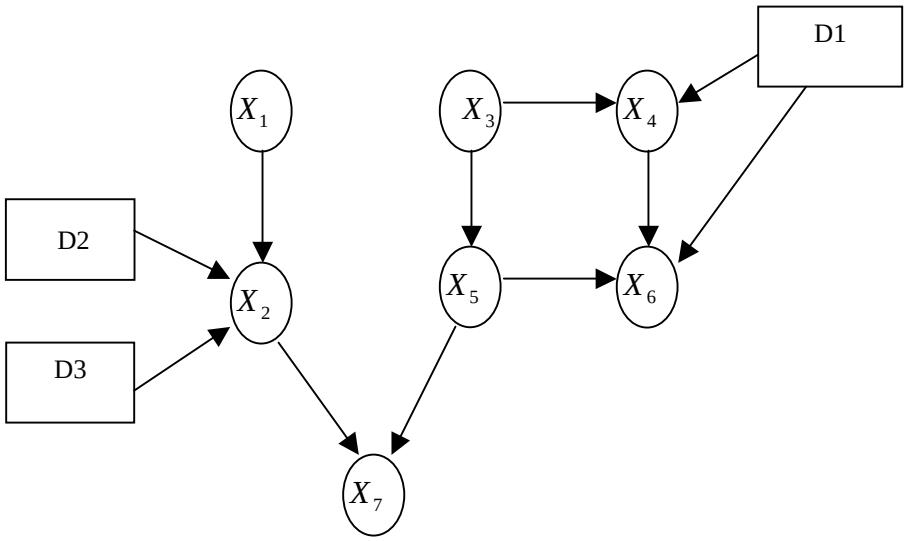


Figure 5 - An Influence Diagram

The HFACS framework, based on Reason's model of latent and active failures [26-27], provides a comprehensive framework to investigate the organization at multiple socio-hierarchical levels. As it originates from Reason's framework, therefore it also inherits the limitations of Reason's model. The HFACS framework qualitatively and quantitatively does not address the interdependencies among different socio-hierarchical elements. A Bayesian Belief Network (BBN), however, provides a probabilistic framework to address and quantify the causal relationships under uncertainty. Fenton, et al. [28] report on the use of BBNs to model the safety assessment of nuclear computer-based systems, for example. By using graph theory, the interrelationships among causal factors can be defined and Bayesian probability theory is used to quantify these relationships. Model quantification occurs by developing Conditional Probability Tables (CPTs) using data when it is available. In the absence of data, an experts' "beliefs" are used. Typically, model quantification involves the fusion of both hard data and "beliefs" and BBNs are ideally suited to such a hybrid or mixed modeling approach. Various elicitation methods of expert beliefs are provided in [29-30]. There are a number of issues concerning human capabilities to consider when eliciting beliefs from the experts as reported in [31-33].

Knowledge elicitation in BBNs involves both qualitative and quantitative forms. Qualitative knowledge comprises identifying causal factors and their interactions by specifying the graphical structure with directed arcs and nodes. Quantitative or semi-quantitative knowledge involves providing numerous conditional probabilities for the BBN. The elicitation of numerous probabilities is typically considered the bottleneck in BBN construction. Renooij [32], Renooij and Witteman [34], van der Gaag, et al. [35], Druzdzel and van der Gaag [36] present an approach to facilitate probability elicitation in BBNs. This approach involves the use fragments of text to provide a conditioning context that are derived from the graphical BBN structure. Then the fragments of text are placed adjacent to a probability scale that contains both verbal probability expressions and numerical values. The verbal expressions are of the form "(almost) certain, probable, expected, fifty-fifty, uncertain, improbable, and (almost) impossible"[32]. The authors contend that the combined approach of both verbal and numerical anchors accelerate conditional probability assessments in BBN when used in conjunction with the fragments of text [32]. This combined approach of using verbal and numerical anchors is being adapted in this

research to assist with probability elicitations for the BBNs dealing with risk assessments of the AvSP technologies/interventions.

The ASRM research uses a case study approach. With a case study approach, statistical generalization is not used. In statistical generalization the samples are chosen randomly and then generalization is observed as a replication of a specific behavior. However, cases are not random samples and each case study represents a unique portrayal. Rather, with case study research, *analytic generalization* and a replication logic is used to generalize to a theory, in this case, system safety theory [18]. Therefore, multiple case studies can be considered as multiple experiments. If two or more case studies show the same behavior, replication can be claimed; however if contrasting results are produced, there should be predictable reasons for this divergent behavior. While specific case studies are used to initiate a dialogue-based process, the resulting influence diagram represents a realistic portrayal of a more generalized model.

Data Analysis: Some key data sources are briefly described below:

HFACS Database: Dr. Scott Shappell (FAA) provided the HFACS data to the Rutgers team. The HFACS data is organized by casual factors such as decision errors, adverse mental states, supervisory violations and organizational processes. It includes all the factors in the generic HUGIN model and also factors such as location and date of the owner, airline owner, etc. Pilots coded the HFACS database that serves as a starting point for the Rutgers research as it helped ground the models in real world data and provided support for establishing connections (i.e. statistical correlations) between causal factors. Figure 6 illustrates the percentage frequency of causal factors for Part 121 scheduled operations. NASA has identified 1990-96 as the baseline period for the AvSP technologies/interventions. The HFACS database may also be used to identify correlations between the different causal factors. Correlation analysis assists with gaining an understanding of the degree to which certain causal factors may be related but it does not necessarily have to be the case for every accident.

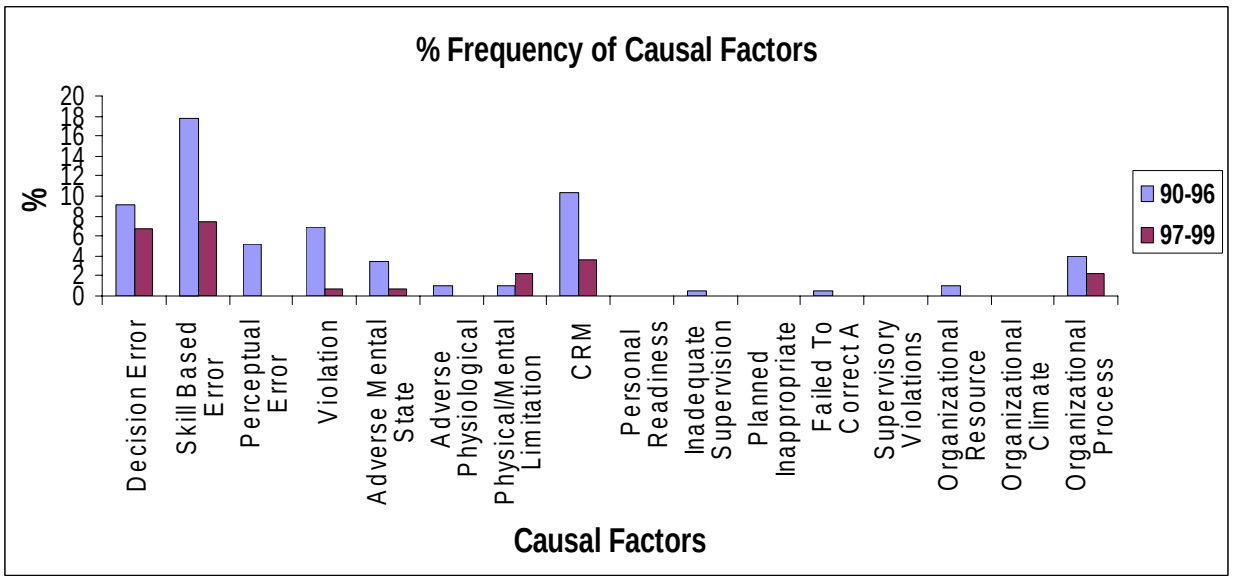


Figure 6 - HFACS Causal Factors

National Aviation Safety Data Analysis Center (NASDAC): The National Aviation Safety Data Analysis Center (NASDAC) [37] is managed by the FAA. The database identifies several causal factors and provides the possibility of identifying different types of accidents. The NASDAC database has a uniform data format and draws on several data sources including the NTSB. Internet accessible, it is possible to specify numerous criteria to limit the results from our data search. The results of the search are displayed in a Microsoft Excel sheet format. It is possible to click on a particular accident report to access a summary description of the final NTSB report for that accident. In this research, the NASDAC database is used to identify LOC cases and to develop an understanding of the overall accident rates for different types of accidents.

Figure 7 graphically portrays the applied approach used in this research. It is a systematic, analytical, dialogue-based approach that initiates with discussions of accident cases with subject matter experts. The causal factors are identified using an expanded HFACS taxonomy, and then the interactions among the causal factors are modeled using influence diagrams. After the influence diagrams are constructed and reviewed by subject matter experts, conditional probability tables are elicited from the "beliefs" or value judgments of the subject matter experts, in conjunction with empirical data where available, to create the BBN. Typically, during this step, 2-3 subject matter experts are used. Generally, a "behavioral aggregation" or consensus-based approach is used during the probability elicitation process, since such an approach encourages the experts to view the final product as a group effort [38]. However, any wide disagreement between the experts is noted for future sensitivity analysis. Then action nodes are added to the BBN to represent the technology/intervention insertions. Additional technologist experts will be included in these discussions. Finally, the projected risk is displayed relative to a baseline period on the relative risk intensity graph.

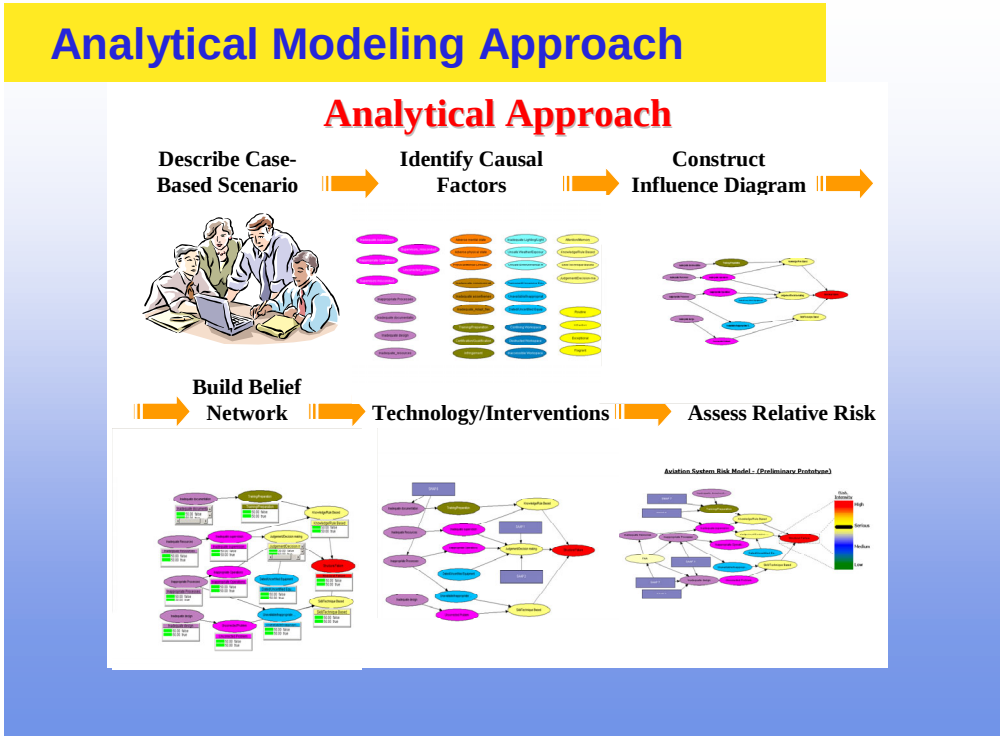


Figure 7 - Applied Research Modeling Approach

Figure 8 graphically illustrates the knowledge acquisition plan for the inclusion of a variety of Safety Inspectors with their general domain knowledge of airworthiness, operations, and avionics are used during the causal modeling and probability elicitation steps. Opportunities will exist for joint discussions between the NASA technologists and the FAA Aviation Safety Inspectors during the technology insertions step. Where possible, the model is supported with data from the NTSB, HFACS, NASDAC, and NASA's ASAFE program. Researchers at the Volpe National Transportation Systems Center are constructing event tree conditional probabilities based on existing aviation safety data sources. These conditional probabilities may be used as "seed" values by the experts that may be modified through the expert elicitation process. It is planned that an Expert Advisory Panel comprised of aviation experts not involved with model construction will be created to review all models and to suggest possible refinements.

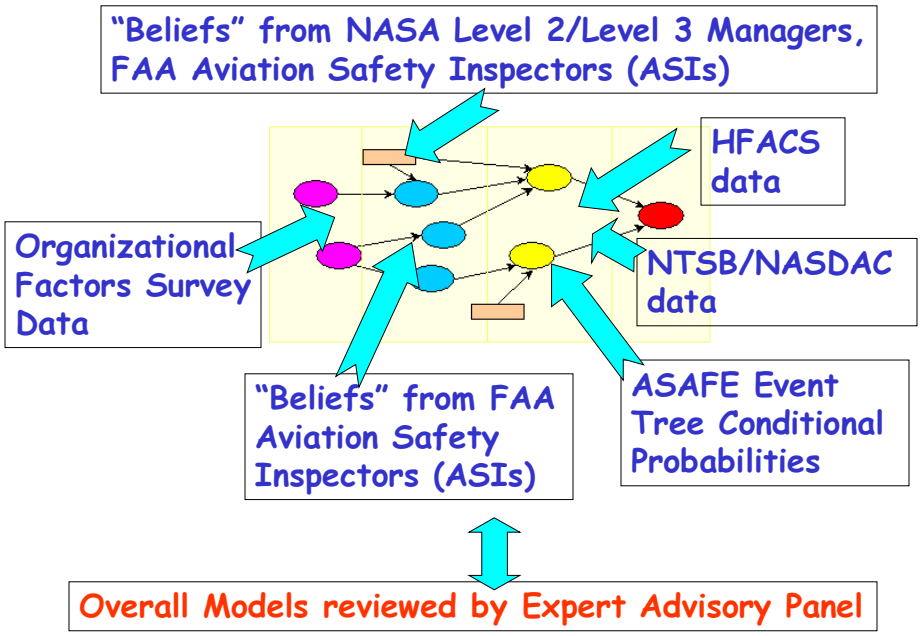


Figure 8 - Multiple Sources Used for "Model Quantification"

"Representative" Loss of Control (LOC) Case: Air Ontario Flight 1363

Accident Summary: The following accident summary presents a "Loss of Control (LOC)" accident where lack of de-icing was a major causal factor. Even though a specific accident is used to provide the "analytical structure" for the case, it should be remembered that through the construct of "analytic generalization" this case is used to generalize to a theory of system safety. The proposed approach is logically consistent while indicating the "causality flows", yet broad enough to consider more general socio-technical factors. The accident summary is adapted from the description and analyses presented in [5].

At 12:09 pm CST on Friday, March 10, 1989, Capt. George C. Morwood, a 24,000-hour flight time veteran, advanced the throttles of Air Ontario Flight 1363, a Fokker F28 1000, initiating the take-off roll at Dryden Airport, Ontario, Canada. Flight 1363 was the second part of the day's

flying schedule that consisted of a Winnipeg to Thunder Bay round trip, with intermediate stops at Dryden (Flights 1362/1363).

Capt. Morwood reviewed the operational status of the aircraft before departing Winnipeg (Flight 1362) and verified, among other maintenance deferred detects, that the Auxiliary Power Unit (APU) was unserviceable. The operational implications of this defect were that the engines had to be started from an external power unit, or one engine had to be kept running to cross-start the other engine. If both engines were shut down at a station where no external power unit was available, the aircraft would be stranded until the APU was fixed or an external power unit became available. There was no external power source at Dryden and therefore one engine would have to be kept running. The manufacturer, Fokker, and Air Ontario strictly prohibited de-icing with either engine running.

Since the original flight release from Thunder Bay to Dryden prepared by the Air Ontario Systems Operations Control (SOC) had not been updated, ten passengers were added to Flight 1363 after it had been refueled. Now overweight for take-off, Capt. Morwood elected to off-load the ten passengers and their baggage. However, the Air Ontario SOC duty manager overrode the captain's decision and chose to achieve weight reduction by off-loading fuel. The de-fueling caused an additional 35-minute delay in the departure of Flight 1363 from Thunder Bay and increased the "hot refueling" time at Dryden. Flight 1363 departed Thunder Bay with a full load of passengers and arrived in Dryden one hour behind the schedule.

The hot refueling process started with passengers on board, which is considered to be an unsafe practice, and is difficult to reconcile with Capt. Morwood's style of decision-making, characterized by conservatism and strict adherence to rules and procedures. The Commission of Inquiry indicated that Capt. Morwood had a heated conversation with the SOC over the telephone regarding the passenger load and weather conditions in Winnipeg prior to the departure from Dryden. The Commission established that the demeanor of Capt. Morwood deteriorated visibly while in the terminal after his telephone contact with the SOC, and that, clearly frustrated, he briskly walked back to the aircraft.

Upon his return to the aircraft Capt. Morwood asked the ground handler whether de-icing was available; however he did not request de-icing after being told that it was. When the aircraft was about to leave the terminal platform, snow was falling heavily, and its wings were covered in snow to depths varying from one-eighth to one-quarter of an inch. While taxiing out, FSS advised Flight 1363 of a Cessna 150 in a VFR recreational flight, which was due to land at Dryden. The pilot of this aircraft had requested that Flight 1363 hold its departure until he had landed, because of the deteriorating weather. The request was eventually granted and the ground hold further compounded the delay of Flight 1363.

The combination of the one-half inch deep slush on the ground and the wet snow, which had frozen into opaque ice on the forward half of the wings, significantly degraded the performance capabilities of the F-28. After a longer than normal take-off roll, the aircraft rotated, lifted off slightly, began to shudder and settled back onto the runway. It rotated a second time, lifting off at the 5,700 ft point of the 6,000 ft runway. It flew briefly, clearing the end of the runway at approximately 15 ft above the ground. It failed to gain altitude and crashed, coming to rest approximately one km. away from the end of the runway.

A Commission of Inquiry was formed on March 29, 1989 to investigate the accident. The Dryden Report, as it has become widely known as, represents one of the first large-scale applications of a systemic, organizational approach to the investigation of an aviation accident.

Aviation System Risk Model (ASRM): The following model depicts the causal factors and the interactions among them.

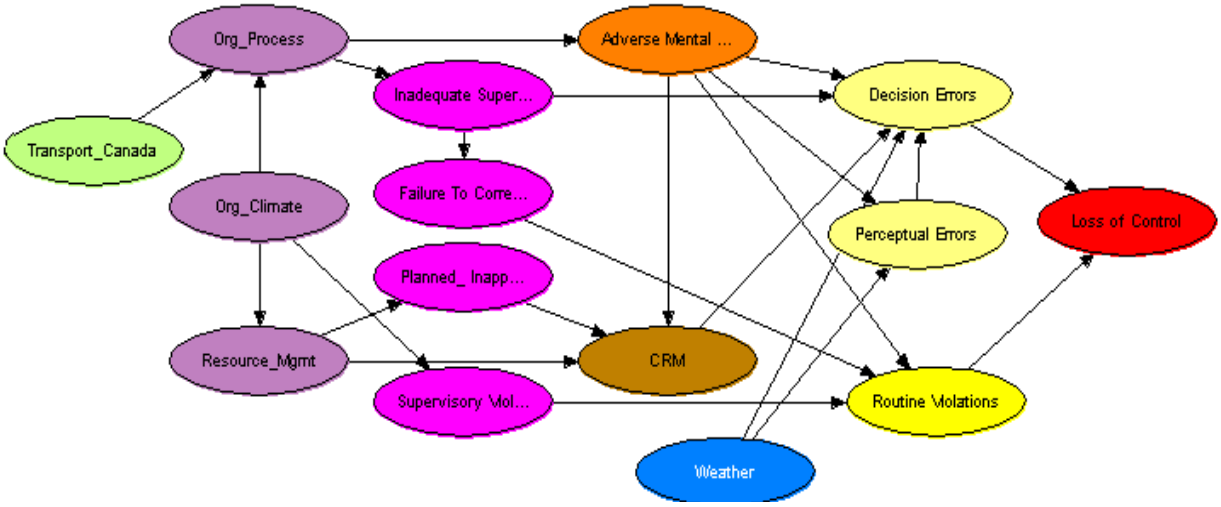


Figure 9 - ASRM for Air Ontario 1363

The interactions among the causal factors of this accident are depicted through the links among the nodes in Figure 9. These interactions were reviewed by the subject matter experts and are based upon a qualitative reasoning process in interpreting the causal factor descriptions provided in the accident report. A snapshot of the reasoning behind a few selected links is provided below:

Organizational Climate ➔ Resource Management
After the merger with Air Canada, Air Ontario was not provided with sufficient levels of training and resources. Due to the inappropriate management style, some of the appointments, such as those of president's close relatives to key managerial positions were the subject of considerable discussions at the Air Ontario committee meetings. Consequently, some Air Ontario managers were confronted by demands for which their experience may not have been adequate.

Inadequate Supervision ➔ Decision Errors
Some F-28 pilots used the Piedmont F-28 Operations Manual while others used the US Air F28 Pilot's Handbook, since Air Ontario did not have its own F-28 operations manual. There was the obvious lack of standardized operations manuals. These deficiencies created problems on the flight deck.

Failed to Correct a Known Problem ➔ Routine Violation
Failures to de-ice and maintenance problems were known by supervisors and yet allowed to continue. Hence, these practices had the form of routine violations.

ASRM with AvSP Technology / Interventions: The definitions and descriptions in this section are based on the NASA Product Dictionary (2002). Based on discussions with subject matter experts and on my research team's knowledge of the 48 technologies, the following represents a

"sample" of the reasoning why certain technologies are added to the model of Air Ontario Flight 1363. Figure 10 depicts the modified ASRM with technology elements.

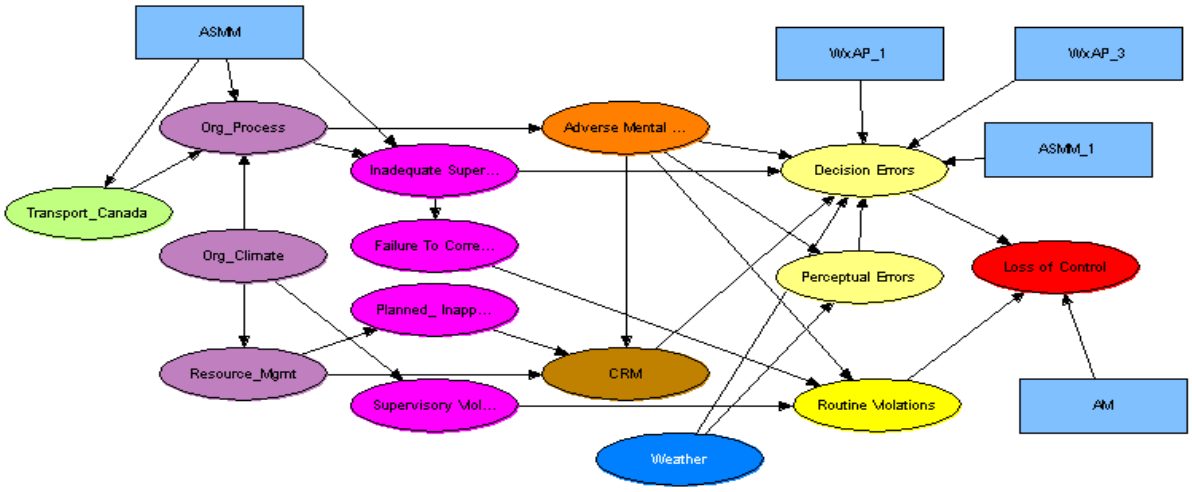


Figure 10 - ASRM for Air Ontario 1363 with Technology Insertions

ASMM4 – Performance Data Analysis & Reporting System (PDARS) Tools: This technology element provides the capability to collect and process Air Traffic Control (ATC) operational data; compute quantitative operational performance measures on a regular basis relating to system safety, delay, flexibility, predictability and user accessibility; conduct causal analyses and operational problem identification and analyses; archive performance statistics and basic operational data for use in research, development and planning studies. The objective here is to monitor performance metrics continuously to enable the implementation of a policy of proactive NAS management. Extending Aviation Performance Measurement System (APMS) concepts to the ATC environment is proposed. Iterative evaluation by ATC users is expected to improve the measurement and tool requirements. Thus, insertion of ASMM4 in 'Transport Canada' node has a potential impact.

WxAP1 – Cockpit Weather System Technologies for Enhanced Situational Awareness & Decision Making: WxAP1 is about the development of substantiated aviation weather information system guidelines for flight deck user interface and for operational use. The objective is to develop enhanced weather presentations that minimize interpretation and enhance situational awareness. Therefore, this technology is considered to have a potential impact on the 'Decision Errors' node.

WxAP3 – Weather Information Datalink Systems Technologies for Ground-to-Air Dissemination: WxAP3 involves the development of datalink system and architecture guidelines supporting the transfer of weather information from ground-to-air. The targeted problem in developing this technology is the poor dissemination of weather information to the flight deck. This technology element is considered to have an impact on the "Decision Errors" node.

Other technology elements included the modified model are:

AM1 – Next Generation Crash Analysis Codes
AM2 – Energy Absorbing Seats, Restraints and Structures

AM4 – Next Generation Crashworthiness Design Guidelines
ASMM1 – Incident Reporting Enhancement Tools
ASMM2 – National Aviation System Operational Monitoring System (NAOMS)

Preliminary Results

Figures 11-12 display screens from the working ASRM prototype that were developed during the Phase 1 research. Figure 12 displays, in general, how the interactions of various causal factors may be depicted, but does not correspond to the aforementioned LOC model. Possible technology insertions are easily portrayed. The working prototype enables sensitivity analyses for both single- and multiple technology insertions. Note that one technology may impact a single and/or multiple causal factors and that multiple technologies may also impact a single and/or multiple causal factors. Changes in relative risk intensity may be displayed on the color-coded bar chart. Both absolute and relative perturbations from a baseline are reported.

Preliminary Risk Assessments: Table 1 illustrates some "representative" preliminary risk assessments using the analytical approach as described previously. The current LOC model has only been through one complete iteration including model quantification with the subject matter experts and thus, any risk assessments should be considered as preliminary. Also, it is intended that an advisory panel will be formed to review all models as well as the beliefs of the experts.

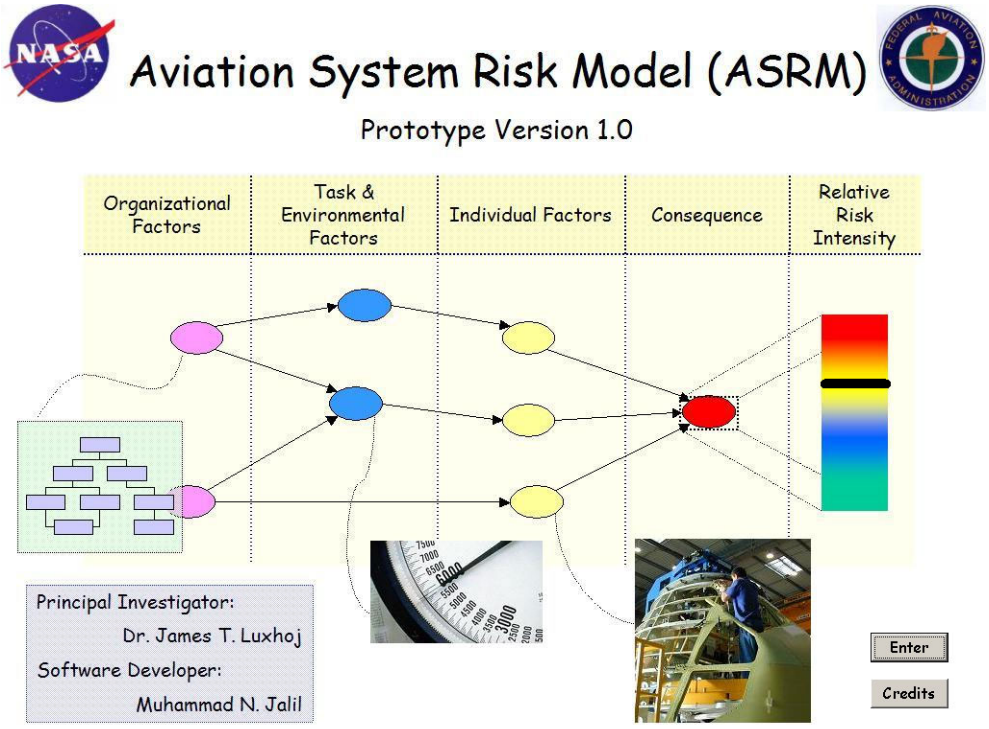


Figure 11 - Initial Screen for the ASRM Prototype

An important point to note from Table 1 is that the overall relative risk reduction on a consequence, such a LOC, may not be as high as the relative risk reduction on a particular causal factor or set of causal factors. The real value of a proposed AvSP technology insertion and/or intervention may lie on the impact on causal factors.

At this point it is important to clarify some terminology used in order to better communicate the analytical approach. A certain modeling "hierarchy" exists, if you will, in the proposed analytical approach. The "Approach" refers to the systematic, step-by-step, ASRM analytical approach as outlined in Figure 7. "Model" is used to refer to the various consequence models, such as Loss of Control (LOC), Controlled Flight Into Terrain (CFIT), Runway Incursion (RI), etc. It should be noted that in Phase 1, four representative "models" are developed for each consequence (LOC1, LOC2,) since there are alternative ways that causal factors could combine for a consequence to occur. Fore example, there could be an LOC accident due to icing or an LOC accident due to improper loading. These models are not exhaustive, but through the construct of *analytic generalization*, we are able to generalize the models to a theory of system safety. For each of the models, various "Scenarios" are created based on alternative combinations of technology insertions/interventions. As noted by Kahn and Wiener [39], a scenario is a "hypothetical sequence of events constructed for the purpose of focusing attention on causal processes and decision-points."

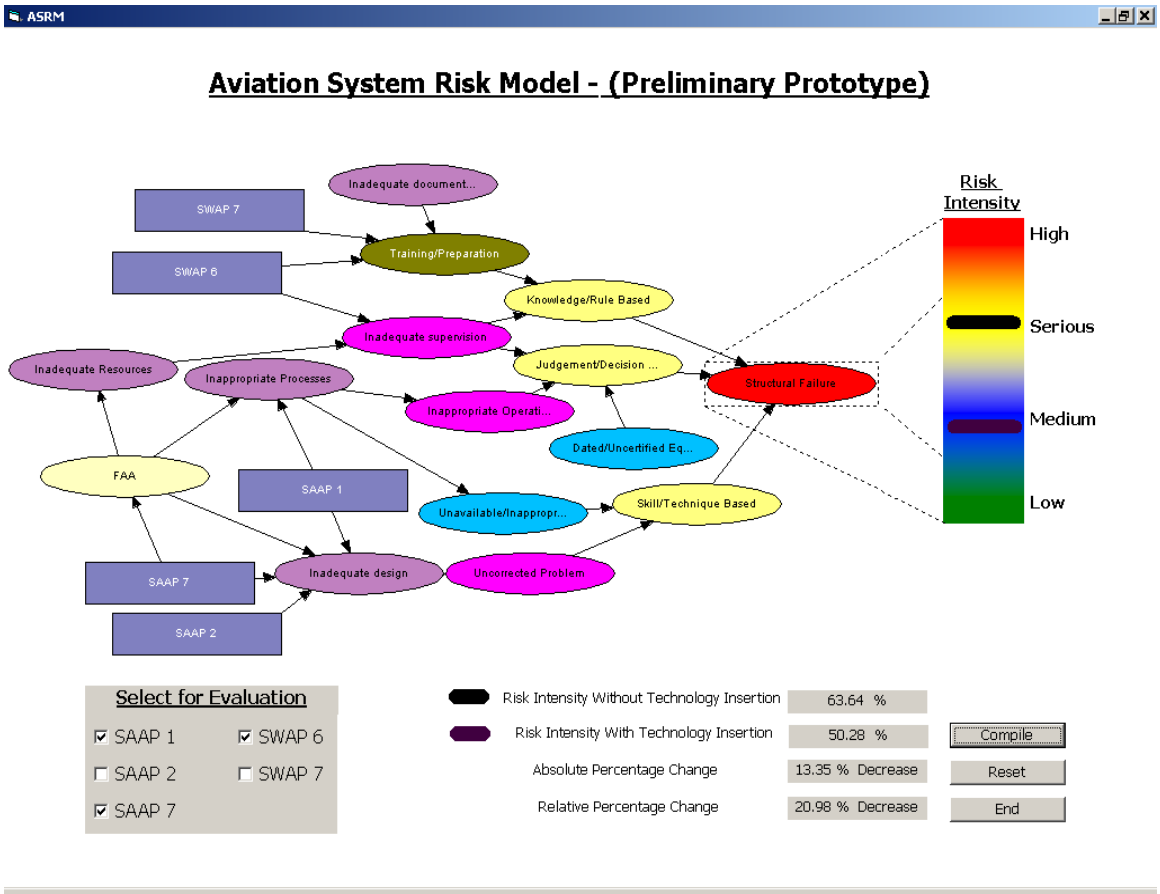


Figure 12 - "Relative Risk Intensity" of the Working ASRM Prototype

In the Phase 1 research, both the LOC and Maintenance accident categories have four models each. Some of the models are not fully quantified at this point in time and some models are partially quantified, so research with these models will continue under Phase 2. In addition, the models need to complete an internal validation step. As previously noted, each of the LOC and Maintenance-related models enables the analysis of *scenario variants*, or different combinations of possible technology insertions. So with the Phase 1 analytical approach, it is possible, for

example, that with 3 models, approximately 30 different scenario variants may be analyzed. This enables both a *literal replication* of cases and a *theoretical replication* of cases [18] and leads to an enrichment of and support for the proposed system safety theoretical approach.

Remaining Research

Over the next three years, the proposed analytical approach and the ASRM software will be used to perform risk assessments for all 48 new technologies/interventions in NASA's AvSP portfolio. The development of these risk assessments will be based on numerous meetings with subject matter experts in the aviation community. In addition, an expert advisory panel will be created to assist with model validation by examining construct validity, internal validity, and external validity as well as repeatability [18]. Belief assessment remains an emerging and developing research field [40].

As an update to the ASRM, it is planned to provide severity distributions for each accident type. As the risk intensity is notionally the likelihood portion of risk, the prototype software will enable the user to drill down to view a "representative" severity distribution and the corresponding risk matrix as illustrated in Figure 13.

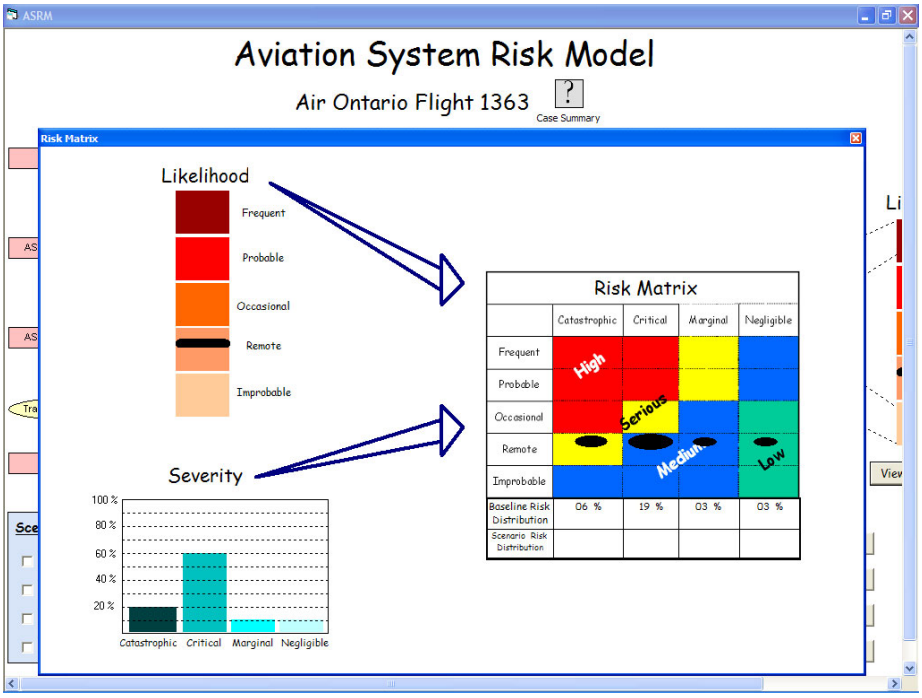


Figure 13 - Inclusion of "Representative" Severity Distribution

Eventually, a suite of models will be developed by accident type, such as Controlled Flight Into Terrain (CFIT), Runway Incursions (RI), etc. Collectively, these models will paint a mosaic of the various contributions that the new technologies/interventions make towards system safety risk reduction in the National Airspace System (NAS).

Acknowledgements

Dr. Luxhøj acknowledges the support of the Federal Aviation Administration (FAA) and NASA through FAA grant number 00-G-006. This research has been extended under NASA contract number NAS1-03057. He also acknowledges the contributions of his research assistants: Mr. Songwut Apirakkhit, Mr. Apichart Choopavang, Mr. Muhammad Jalil, Mr. Erim Kardes, Ms. Kimberlee Kauffeld, Mr. Ram Kuturu, Mr. Ahmet Oztekin, Mr. Ryan Dickey, Mr. Nathan Greenhut, Ms. Denise Andres, as well as Dr. David Coit. Dr. Luxhøj and his research team are also grateful to the FAA's Aviation Safety Inspectors who served as initial subject matter experts for this research. This paper is based on the research performed at Rutgers University. The contents of this paper reflect the views of the author who is solely responsible for the accuracy of the facts, analyses, conclusions, and recommendations represented herein, and do not necessary reflect the official view or policy of either the Federal Aviation Administration or NASA.

Table 1 - LOC 1 (ICING) – Air Ontario 1363

Description	Scenario	Targeted causal factor(s)	Technology element(s) inserted	Risk intensity (Factors)	Absolute % Decrease or (Increase) On factors	Relative % Decrease or (Increase) On factors	Risk Intensity (Consequence)	Absolute % Decrease or (Increase) on consequence	Relative % Decrease or (Increase) on consequence
Baseline scenario	No technology insertion	----	----	----	-	-	31.14% (Medium)	-	-
LOC 1 Scenario 1	ASMM suite	Regulator Org. Process Res. Mgmt. CRM	ASMM 1,2,3,4,5,6	N/A	N/A	N/A	28.34% (Medium)	2.8%	8.99%
LOC 1 Scenario 4	WxAP suite	Org. Process Decision Errors	WxAP 1,2,3,4	N/A	N/A	N/A	26.76% (Medium)	4.38%	14.06%
LOC 1 Scenario 5	Effect of interventions on regulator	Regulator	SAAP-4 ASMM 1,2,3,4,5,6	18.00%	4.00%	22.22%	28.14% (Medium)	3.00%	9.63%
LOC 1 Scenario 7	Effect of interventions on Decision Errors	Decision Errors	WxAP 1,2,3,4	34.27%	8.72%	25.44%	26.76% (Medium)	4.38%	14.06%
LOC 1 Scenario 8	Effect of interventions on Org. Process	Org. Process	WxAP 2, SAAP 4, ASMM 1,2,3,6	69.09%	15.89%	22.99%	28.00% (Medium)	3.14%	10.08%
LOC 1 Scenario 10	Effect of all possible interventions	Model	ASMM 1,2,3,4,5,6 SWAP 1,2 SAAP 4 WxAP 1,2,3,4	N/A	N/A	N/A	23.75% (Medium)	7.39%	23.73%

References

1. Leveson, N., "A New Foundation for System Safety" (http://sunnyday.mit.edu), 2002.

2. Leplat, J., "Occupational Accident Research and Systems Approach," in *New Technology and Human Error* (J. Rasmussen, K. Duncan, and J. Leplat, eds.) New York: John Wiley and Sons, Inc., 1987, pp. 181-191.

3. Pate-Cornell, M.E., "Organizational Aspects of Engineering System Safety: The Case of Offshore Platforms," *Science* 250, 1990, pp. 1210-1217.

4. Pidgeon, N. and M. O'Leary, "Organizational Safety Culture: Implications for Aviation Practice," in *Aviation Psychology in Practice* (N. Johnston, N. McDonald, and R. Fuller, eds.), Hove, The Netherlands: Lawrence Erlbaum, 1994.

5. Maurino, D.E., J. Reason, N. Johnston, and R.B. Lee, *Beyond Aviation Human Factors*, Vermont: Ashgate Publishing Company, 1997.

6. Reason, J., Managing *the Risks of Organizational Accidents.* England: Ashgate Publishing Limited, 1997.

7. Sarsfield, L.P., "Aviation in the Next Millennium: A National R&D Plan for Improving Air Transportation Safety and Security in the 21st Century", Draft, RAND Science and Technology Policy Institute, October 1, 1998.

8. Strauch, B., *Investigating Human Error: Incidents, Accidents, and Complex Systems*, England: Ashgate Publishing Limited, 2002.

9. NASA Aviation Safety Program (AvSP) Product Dictionary, February 2002.

10. NASA Aviation Safety (AvSP) Program Commitment Agreement, July, 2000.

11. Luxhøj, J.T., D.N. Arendt, T.P. Williams, and T.G. Horton, "An Application of Advanced Information Technology for Assessing Aircraft Accident Causation," Proceedings *of International Society of Air Safety Investigators*, Anchorage, Alaska, September 29 - October 3, 1997.

12. Luxhøj, J.T., D.N. Arendt, and T.G. Horton, "An Intelligent Computer-Based Tool for Evaluating Aircraft Accident Causation," *Proceedings of European Safety and Reliability International Conference ESREL '98*, Trondheim, Norway, June 17-19, 1998, pp. 839-846.

13. Luxhøj, J.T., K. Bansal, A. Choopavang, A., T.G. Horton, and D. N. Arendt, "An Aviation System Safety Model for Improved Risk and Management," *Proceedings of the European Safety and Reliability Conference ESREL'99*, Munich, Germany, September 13-17, 1999, pp. 1285-1290.

14. Luxhøj, J.T., A. Choopavang, and D.N. Arendt, "Risk Assessment of Organizational Factors in Aviation Systems," *Air Traffic Control Quarterly*, 9 (3), *(Special Issue on Flight Safety)*, 2001, pp. 135-174.

15. Andersen, S.K., F.V. Jensen, and K.G. Olesen, "The HUGIN Core – Preliminary Considerations on Systems for Fast Manipulation of Probabilities," in *Proceedings of Workshop on Inductive Reasoning: Managing Empirical Information in AI-Systems,* Risø National Laboratory, Roskilde, Denmark, April, 1987.

16. Andersen, S.K., K.G. Olesen, F.V. Jensen, and F. Jensen , "HUGIN - A Shell for Building Bayesian Belief Universes for Expert Systems," in *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, Detroit, Michigan, August 20-25, 1989, pp. 1080-1085.

17. Jensen, F.V., S.L. Lauritzen, and K.G. Olesen, "Bayesian Updating in Causal Probabilistic Networks by Local Computations," *Computational Statistics Quarterly,* 4, 1990, pp. 327-352.

18. Yin, R., *Case Study Research: Design and Methods, 3rd ed.*, London: Sage Publications, 2003.

19. van Vuuren, W., Organizational *Failure: An Exploratory Study in the Steel Industry and Medical Domain*, Institute of Business Engineering and Technology Application, Eindhoven University of Technology, Eindhoven, Netherlands, 1998.

20. Shappell, S.A., and D.A. Wiegmann, "The Human Factors Analysis and Classification System-HFACS," Office of Aviation Medicine Technical Report No, DOT/FAA/AM-00/7, Civil Aeromedical Institute, Oklahoma City, OK, February, 2000.

21. Shappell, S.A. and D.A. Wiegmann, "Applying Reason: The Human Factors Analysis and Classification System (HFACS)," *Human Factors and Aerospace Safety*, 1 (1), 2001, pp. 59-86.

22. Wiegmann, D.A., and S.A. Shappell, "A Human Error Analysis of Commercial Aviation Accidents Using the Human Factors Analysis and Classification System (HFACS)," Report Number DOT/FAA/AM-01/3, Washington DC: Federal Aviation Administration, 2001.

23. Gardiner, P., and R. Wood, "Operational Risk Management," *Handbook of Airline Operations,* New York: McGraw-Hill, 2000.

24. Pearl, J., *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, San Francisco: Morgan Kaufmann, 1988.

25. Jensen, F.V., *Introduction to Bayesian Networks*, New York: Springer-Verlag, 1996.

26. Reason, J., *Human Error*, Cambridge, UK: Cambridge University Press, 1990.

27. Reason, J., "A System Approach to Organizational Error," *Ergonomics* 38(8), 7, 1995, pp. 1708-1721.

28. Fenton, N.E., B. Littlewood, M. Neil, L. Strigini, D.R. Wright, and P.J. Courtois, "Bayesian Belief Network Model for the Safety Assessment of Nuclear Computer-based Systems, ESPRIT DeVa Project 20072, Center for Software Reliability, City University, London, January, 1998.

29. Ayyub, B.M., *Elicitation of Expert Opinions for Uncertainty and Risks,* New York: CRC Press, 2001.

30. Vick, S., *Degrees of Belief: Subjective Probability and Engineering Judgment,* Reston, VA: ASCE Press, 2002.

31. Wilson, A.G., "Cognitive Factors Affecting Subjective Probability Assessment," *ISDS Discussion Paper # 94-02*, Institute of Statistics and Decision Sciences, Duke University, February, 1994.

32. Renooij, S., "Probability Elicitation for Belief Networks: Issues to Consider," *The Knowledge Engineering Review,* 16(3), 2001, pp. 255-269.

33. Gilovich, T., D. Griffin, and D. Kahneman, *Heuristics and Biases: The Psychology of Intuitive Judgment,* New York: Cambridge University Press, 2002.

34. Renooij, S. and C.L.M. Witteman, "Talking Probabilities: Communicating Probabilistic Information With Words and Numbers," *International Journal of Approximate Reasoning,* 22, 1999, pp. 169-194.

35. van der Gaag, L.C., S. Rennoij, C.L.M. Witteman, B.M.P. Aleman, and B.G. Taal, "How to Elicit Many Probabilities," *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence,*1999, pp. 647-654.

36. Druzdzel, M.J. and L.C. van der Gaag, "Elicitation of Probabilities for Belief Networks: Combining Qualitative and Quantitative Information," *IEEE Transactions on Knowledge and Data Engineering,* 12(4), 2000. pp. 481-486.

37. *NASDAC Database Report No. AAR 96*-06, National Aviation Safety Data Analysis Center, Federal Aviation Administration, Washington DC, 1996.

38. Meyer, M.A., and J.M. Booker, *Eliciting and Analyzing Expert Judgment: A Practical Guide,* New York: Academic Press, 1991.

39. Kahn, W. and A.J. Wiener, *The Year 2000: A Framework for Speculation on the Next Thirty-Three Years*, New York: Macmillan, 1968.

40. Benson, P.G., S.P. Curley, and G.F. Smith, "Belief Assessment: An Underdeveloped Phase of Probability Elicitation," *Management Science,* Vol. 41, No. 10, October 1995, pp. 1639-1653.

Biography

J.T. Luxhøj, Rutgers University, Department of Industrial and Systems Engineering, Piscataway, NJ 08854-8018 U.S.A.; telephone 732.445.3625; fax 732.445.5467; e-mail – jluxhoj@rci.rutgers.edu

Dr. James T. Luxhøj is Associate Professor of Industrial and Systems Engineering at Rutgers University. In 1994-95 and Fall 2001 he was a Visiting Professor at Aalborg University in Denmark. He received his Ph.D. in industrial engineering and operations research from Virginia Polytechnic Institute and State University in 1986. He has been involved in aviation systems analysis over the past 12 years. He served as the Principal Investigator on Federal Aviation Administration (FAA) research grants to develop an intelligent decision support system for aviation safety analysis and to develop analytical methods for aviation safety risk modeling, assessment, and management. Jim is currently collaborating with NASA's Aviation Safety

Program to apply risk modeling techniques to evaluate the impact of technology insertion upon system risk in the National Airspace System. He is the former Co-Chair for the international GAIN Working Group B: Analytical Methods and Tools and has served as the Co-Chair of the FAA's recent National Workshops on Risk Analysis and Safety Performance Measurement in Aviation. He has published extensively on topics such as risk analysis, reliability and maintenance modeling, econometric modeling, and decision support systems. Dr. Luxhøj serves as a Department Editor for *IIE Transactions on Operations Engineering*, and as Associate Editors for the *Journal of Design and Manufacturing Automation* and the *Journal of Engineering Valuation and Cost Analysis*. Jim resides in Somerset, New Jersey with his wife and two children.

Risk-based Classification of Incidents

William S. Greenwell, John C. Knight, Elisabeth A. Strunk,
Department of Computer Science, University of Virginia; Charlottesville, VA, U.S.A.

Keywords: safety-critical systems, incident investigation, risk assessment

Abstract

As the penetration of software into safety-critical systems progresses, accidents and incidents involving software will inevitably become more frequent. Identifying lessons from these occurrences and applying them to existing and future systems is essential if recurrences are to be prevented. Unfortunately, investigative agencies do not have the resources to fully investigate every incident under their jurisdictions and domains of expertise and thus must prioritize certain occurrences when allocating investigative resources. In the aviation community, most investigative agencies prioritize occurrences based on the severity of their associated losses, allocating more resources to accidents resulting in injury to passengers or extensive aircraft damage. We argue that this scheme is inappropriate because it undervalues incidents whose recurrence could have a high potential for loss while overvaluing fairly straightforward accidents involving accepted risks. We then suggest a new strategy for prioritizing occurrences based on the risk arising from incident recurrence.

Introduction

By their very nature, commercial aviation accidents demand our attention. Major accidents can create spectacular scenes of carnage and destruction that threaten public confidence in commercial air travel. At the very least, accidents remind us that, while very safe, there is still some risk in commercial air travel, and they often force engineers and regulators to rethink their safety analyses and add additional safeguards to the air transit system. It is out of a desire to improve safety and prevent the recurrence of tragedy that society demands investigations into accidents in order to learn as many lessons from them as possible.

Although major accidents receive the most publicity, less severe accidents and even incidents in which no loss is incurred can be equally valuable in their ability to provide lessons [2]. Despite this, incidents rarely command the attention that accidents do, and this is a serious imbalance with possibly serious consequences. This paper presents two commercial aviation events involving safety-critical software systems in which the failure of those systems contributed to the occurrence of the events. The first event resulted in a crash with hundreds of fatalities. Although the second event did not develop into an accident, the failure of the system involved led to a near-collision between two jumbo jets. After summarizing the events, we show that the first event received a much more rigorous investigation than the second, even though the latter could have resulted in almost twice the number of fatalities. We then suggest an alternative incident classification scheme that we claim will more appropriately match investigative resources to events whose recurrence would likely have catastrophic consequences.

Both of the events that we discuss in this paper could have been prevented in many ways. However, the need for change in incident classification is illustrated very clearly by the fact that both events were *preceded* by similar incidents that indicated the possibility of a systemic problem [3]. Our strategy attempts to exploit such leading indicators to prevent future accidents.

Accidents Versus Incidents

Before proceeding, it is useful to distinguish accidents from incidents. Numerous definitions exist for these terms; however because this paper focuses on two commercial aviation events, the definitions we use are those adopted by the Federal Aviation Administration (FAA) and the National Transportation Safety Board (NTSB). Those organizations define the terms as follows:

> *Aircraft accident*—an occurrence associated with the operation of an aircraft which takes place between the time any person boards the aircraft with the intention of flight and all such persons have disembarked, and in which any person suffers death or serious injury, or in which the aircraft receives substantial damage.
>
> *Incident*—an occurrence other than an accident, associated with the operation of an aircraft, which affects or could affect the safety of operations [11].

According to these definitions, for an unsafe occurrence involving an aircraft to be considered an accident, it must take place when persons are aboard the aircraft and result in some form of loss—death, serious injury, or damage to the aircraft. Otherwise, the occurrence is considered to be an incident. Most aircraft accidents and incidents occur while the aircraft is in operation, which implies that persons will be aboard at the time of an occurrence. Thus, loss is the key factor that distinguishes an accident from an incident, which agrees with the distinction made by Leveson [16]. All occurrences affecting safety begin as incidents, and whether they remain incidents or develop into accidents depends upon their outcomes. Incidents and accidents might share similar event sequences, meaning that incidents are sometimes precursors to accidents. Investigating incidents can lead to recommendations that help prevent future accidents.

Accident & Incident Investigation

Unfortunately, most investigative agencies simply do not have the resources to fully investigate every aviation-related occurrence within their jurisdiction and must prioritize certain occurrences when allocating investigative resources. Agencies typically prioritize occurrences according to the severity of their associated losses. For example, the NTSB classifies an accident as “major” if the accident results in the destruction of a commercial aircraft, multiple fatalities, or one fatality and substantial damage to a commercial aircraft. According to NTSB statistics, 74 major accidents occurred between 1983-2002 compared to 581 accidents receiving less severe designations involving commercial aircraft [5]. NTSB investigators use a special operating manual when investigating major accidents that guides them in collecting evidence, holding public hearings, and preparing final reports [6]. Reports are typically reserved for major accidents; synopses are prepared for less severe accidents and then stored in a database. The NTSB investigates all civil aviation accidents that occur within its jurisdiction. It also selectively investigates aviation incidents but is not required to do so. NTSB incident reports are stored in the board’s accident database and resemble the synopses it prepares for minor accidents.

In addition to the NTSB, the FAA may also investigate civil aviation incidents at its discretion, and informal channels exist for collecting and analyzing incident reports such as the Aviation Safety Reporting System (ASRS), which is administered by the FAA and the National Aeronautics and Space Administration (NASA). Pilots, air traffic controllers, flight attendants, mechanics, and others may voluntarily submit incident reports to the ASRS. These reports are reviewed by ASRS personnel who use them to prepare monthly safety bulletins and to identify immediate safety hazards to report to the FAA [12]. A similar system called CHIRP exists in the United Kingdom. The FAA also maintains partnerships with air carriers and repair stations

known as Aviation Safety Action Programs (ASAPs) that encourage employees to voluntarily report information that might help the FAA identify potential precursors to accidents [15]. Johnson notes, however, that these systems have limitations and in particular tend to focus on direct, short-term fixes to safety problems rather than addressing underlying issues [13].

The NTSB is not unique in prioritizing accident and incident investigations according to loss. Most investigative agencies worldwide distinguish between accidents and "serious incidents," including the U.K. Air Accidents Investigation Branch (AAIB), the French Bureau d'Enquêtes et d'Analyses (BEA), the German Federal Bureau of Aircraft Accidents Investigation (BFU), the Accident Investigation Board of Finland (AIB), the Australian Transport Safety Bureau (ATSB), the Taiwanese Aviation Safety Council (ASC), and others. While the definition of "accident" is typically clear, the term "serious incident" is often not well-defined. The AAIB and BFU offer guidelines that give examples of serious incidents, but admit that these guidelines are not comprehensive. The ATSB uses a five-category system to classify accidents and incidents, but the criteria for categorizing an occurrence are subjective. The Canadian Transport Safety Board (TSB) does not actually distinguish between accidents and incidents but labels both types of events as "occurrences." They classify and investigate occurrences based on "whether the investigation is likely to lead to reduced risk to persons, properly, or the environment" [9]. This is similar to the scheme we propose; however their criteria are still quite subjective.

The effect of allocating resources to accident and incident investigations based on the severity of their associated losses is that less severe accidents might receive only a small amount of attention from investigators, and incidents might not be investigated at all. However, many major *accidents* are preceded by similar *incidents* in which it was only by coincidence that a loss did not occur. This is particularly important in the context of safety-critical software systems because design faults present in such systems can manifest themselves with unpredictable consequences. If the systems control hazardous operations, they might bring direct harm to passengers or crew. Alternatively, if the systems provide advice or warnings to pilots, they might raise false alerts or issue erroneous guidance to pilots, who could inadvertently jeopardize safety by acting on this information.

To illustrate the disparity in the level of attention given to accidents versus that typically given to incidents, we examine the investigations conducted following a major accident and a major incident. The following sections begin with brief descriptions of the occurrences and then present details of their respective investigations.

Korean Air Flight 801

On August 6, 1997 at about 1:42am Guam local time, Korean Air flight 801, a Boeing 747-300, crashed into Nimitz Hill, Guam while attempting a nonprecision approach to runway 6L at A.B. Won Guam International Airport. Of the 254 persons on board, 237 of which were passengers, only 23 passengers and 3 flight attendants survived. The National Transportation Safety Board (NTSB) investigated the accident and classified the crash as a controlled-flight-into-terrain, or CFIT, accident. During its investigation, the NTSB found that a ground-based minimum safe altitude warning system (MSAW), designed to alert air traffic controllers of aircraft flying too low, had been inhibited. In its final report [7], the NTSB concluded that the crash was largely due to pilot error, but also noted:

> "Contributing to the accident was the Federal Aviation Administration's (FAA) intentional inhibition of the minimum safe altitude warning system (MSAW) at Guam and the agency's failure to adequately manage the system."

We discuss in detail how the MSAW system at Guam contributed to the accident elsewhere [3]. Essentially, the system had been disabled years before the accident in order to eliminate nuisance low-altitude warnings. Prior to the accident, the FAA had received multiple warnings that MSAW systems were being configured improperly. These included a safety recommendation from the NTSB issued in response to a previous accident urging the FAA to verify the MSAW configurations at each of its air traffic control facilities as well as an evaluation of the Guam facility that noted its MSAW inhibition. After the Korean Air flight 801 accident, the FAA developed a comprehensive program to manage its MSAW installations, but continued to be plagued by accidents in which MSAW configuration errors were cited as contributory factors.

The NTSB began its investigation into the Korean Air flight 801 accident immediately after the crash. The Board adopted its final report, a 212-page document, on January 13, 2000. The report contains 134 pages of factual information pertaining to the accident and 37 pages of analysis. The investigation yielded 36 findings and a set of 15 recommendations mostly addressed to the FAA. During the investigation, the NTSB held a three-day public hearing into the accident in which officials from the FAA, Korean Air, the government of Guam, and other organizations gave testimony. The transcript from this hearing spans approximately 430 pages [8].

British Airways Flight 027

On June 28, 1999, British Airways flight 027, a Boeing 747 carrying 419 passengers and crew members en route to Hong Kong, China, and another Boeing 747 operated by Korean Air Cargo nearly collided in flight over a remote region of Chinese airspace. At their closest point of approach, the two aircraft passed within 600 feet of each other, and the British Airways copilot later recounted that his windshield was consumed by the fuselage of the other jet. No injuries resulted from the incident and both aircraft arrived at their destinations. If the two aircraft had collided, however, it is likely that none of the persons aboard either aircraft would have survived [10].

Prior to the incident, the two aircraft were travelling in opposite directions along the same airway with a safe margin of 2,000 feet of vertical separation. The British Airways passenger flight was flying above the Korean Air Cargo flight. The incident sequence began when a collision avoidance system onboard the Korean Air Cargo flight malfunctioned and mistakenly determined the aircraft's altitude to be 2,400 feet higher than its true altitude. This caused the system to believe that a traffic conflict existed between the two aircraft, which prompted it to erroneously instruct the Korean Air pilot to climb in order to avoid the conflict. Because no air traffic control service was available in the region of airspace in which the aircraft were operating and meteorological conditions prevented the pilots from visually identifying each other's aircraft, the Korean Air pilot had no reason to question the collision avoidance system's instruction and thus complied. This placed the two aircraft on a collision course that neither flight crew detected until moments before the aircraft reached their closest point of approach. British Airways officials later noted that it was only by coincidence that the two aircraft avoided each other and that they would have likely collided had they been using more precise navigation systems such as the Global Positioning System (GPS) navigation systems in widespread use today [1].

With the assistance of Korean Air, the CAA determined that the malfunction in the Korean Air Cargo jet's collision avoidance system was caused by damage inflicted during maintenance to the aircraft's avionics systems. Upon concluding its investigation, the CAA issued an airworthiness directive requiring air carriers using similar systems to periodically conduct inspections to ensure the systems are using correct altitude values. The CAA also notified other European aviation regulatory agencies, the FAA, and equipment manufacturers of the problems it found, and it

issued a recommendation to aircraft operators urging them to consider using more robust schemes for handling altitude data.

The U.K. Civil Aviation Authority (CAA) and British Airways each conducted their own investigations into the incident. The CAA's report does not indicate when its investigation into the incident began; however the report is dated October 28, 1999, suggesting that the investigation lasted at most four months. The report is three pages long and includes eight paragraphs of factual information spanning two pages and a single paragraph of analysis. It contains a single conclusion and three recommendations directed at operators and equipment manufacturers. No public hearing was held in response to this incident. British Airways prepared a more detailed report on the incident, but that report has not been officially released to the public.

Event Comparison

In order to help quantify the difference in rigor for the investigations described earlier, we have summarized data from the events and their investigations in Table 1 below.

Table 1 - Comparison of Korean Air flight 801 and British Airways flight 027

	Korean Air 801	British Airways 027
Classification	Accident	Incident
Persons On Board	254	419
Fatalities	228	0
Injuries, Serious	26	0
Injuries, Minor	0	0
Total Casualties	254	0
Aircraft Damage	Destroyed	None
Investigation Length (months)	30	4
Final Report Length (pages)	212	3
Factual Information (pages)	134	2
Analysis (pages)	37	1
Findings / Conclusions	36	1
Recommendations	15	3

The first seven fields listed in Table 1 assess the loss from each incident and the remaining fields attempt to capture the level of rigor applied in the subsequent investigations. Examining the fields pertaining to loss, the near-collision involving British Airways 027 had no casualties compared to

a 90% fatality rate in the Korean Air 801 accident. In addition, neither of the Boeing 747s involved in the near-collision sustained any damage from the incident, whereas the 747 involved in the Guam accident was destroyed.

Comparing these events solely on the basis of loss is deceiving, however, as the British Airways incident could have easily developed into an accident with almost twice the number of fatalities as the Korean Air flight 801 crash in Guam. As British Airways officials noted, it was entirely by luck that the British Airways passenger flight and the Korean Air Cargo flight did not collide. By the time the Korean Air pilot inadvertently placed his aircraft on a collision course with British Airways flight 027, all of the barriers designed to prevent midair collisions had been defeated, and conditions were sufficient for a collision to occur. Indeed, if the incident were repeated under similar circumstances it is likely that a collision would occur, which suggests that the risk of a recurrence of the British Airways flight 027 incident is at least as severe as that of a recurrence of the Korean Air flight 801 accident if not more so.

Given the risk of a recurrence of the near-collision, one would expect a thorough investigation to be conducted in order to determine what prompted the Korean Air Cargo pilot to suddenly veer toward the aircraft flying above. The remaining fields in Table 1 suggest that this was not the case. While in general criteria such as investigation length, report length, and number of findings or recommendations are not indicative of an investigation's thoroughness, the differences indicated in Table 1 between the two incidents are too extreme to ignore. The factual information, analysis, findings, and recommendations from the CAA's investigation into British Airways flight 027 are only a fraction of those from the NTSB's investigation into Korean Air flight 801. This is not because the former was a simple incident. On the contrary, several factors contributed to the loss of separation and subsequent near-collision, including design faults present in the incident aircraft's collision avoidance systems and the systems they interfaced with, human factors issues concerning the manner in which traffic information was displayed to the flight crews, and broader issues concerning the role that collision avoidance systems play in the overall air traffic system. The CAA's report failed to examine these issues, and consequently missed an opportunity to correct problems that might contribute to future incidents, possibly with more dire outcomes.

Under the loss-based accident classification schemes employed by most investigative agencies, such a catastrophic outcome would be necessary for a major investigation to be undertaken, even though the findings and recommendations would likely be the same as if an equally rigorous investigation had been conducted into the incident alone. This should not be the case. New classification schemes are necessary in order to better allocate investigative resources to incidents whose recurrence could have more severe consequences.

In reviewing this comparison, one might argue that the vast difference between the Korean Air and British Airways events was not necessarily because of their associated losses but rather due to the fact that different agencies investigated each event. Had both events been investigated by the NTSB or CAA, the figures might have matched more closely. The NTSB's incident reports tend to match the CAA's report in length, however, and often do not contain immediate safety recommendations (although incident data is aggregated for use in later recommendations). Similarly, if the Korean Air flight 801 accident had occurred in British airspace, it would have been investigated not by the CAA but by the AAIB, whose formal reports are similar to the NTSB's final reports in structure and length.

Risk-based Classification of Incidents

The term "incident" can be defined in a variety of ways but typically involves the failure of a network of barriers designed to protect a system from one or more hazards. An incident becomes an accident when it is coupled with a loss event such as a crash or collision in which damage or casualties are incurred. It is often the case that luck determines whether an incident develops into an accident and, if so, what the extent of the loss will be.

When investigating accidents, investigators can issue recommendations aimed at preventing the associated incident or at mitigating the severity of the loss, and they usually do both. While attempting to mitigate loss given the occurrence of an incident can help to reduce the severity of accidents, some degree of loss is almost always inevitable. On the other hand, if the incident itself is prevented, it cannot develop into an accident and thus no loss will occur. Therefore, recommendations aimed at preventing incident recurrences are likely to be more effective in preventing future losses. Indeed, 13 of the 15 recommendations issued by the NTSB in response to the Korean Air flight 801 accident were aimed at preventing the recurrence of incidents in which aircraft descend below safe altitudes during final approach. Only two focused on mitigating losses by suggesting improvements to Guam's emergency response units.

Given that accidents begin as incidents and that incident prevention should be the focus of investigations, incidents are opportunities for investigators to identify problems and suggest safety improvements without the losses associated with accidents. Accident classification schemes based on loss alone place a low priority on incidents even though those incidents might be indicative of safety problems that could lead to more catastrophic outcomes should they recur. By itself, loss is a poor indicator of an incident's potential for learning new lessons and preventing future incidents. Therefore, classification schemes based on loss should be de-emphasized in favor of new schemes in which resources are allocated to incident investigations based on the risk associated with the incidents' recurrence. To this end, the fundamentals for such a scheme are presented below.

Risk is defined as the probability that an event will occur multiplied by the anticipated cost derived from the occurrence of the event. When an incident occurs, it suggests the presence of a deficiency in the safety systems involved that, if not corrected, could lead to recurrences of the incident. A useful measure of the importance of an incident, therefore, is the total risk that society faces if nothing is done to prevent recurrences. The total risk of such a recurrence is given in Equation 1 below.

$$\begin{aligned} \text{Total Risk} &= E[\#\text{ Recurrences}] \times E[\text{Cost}] \\ &= P[\text{Incident Recurrence}] \times \text{Exposure} \times E[\text{Cost}] \end{aligned} \tag{1}$$

The term E[# Recurrences] represents the expected number of recurrences of the incident if nothing is done to reduce the likelihood of recurrence and is the product of P[Incident Recurrence], the probability that the incident will happen again, and exposure, the number of opportunities for the incident to recur. The term E[Cost] is the expected cost of the incident given that it has occurred and is defined in Equation 2 below.

$$E[\text{Cost}] = \sum_{i \in S} \text{Cost}(i) \bullet P[i] \tag{2}$$

Equation 2 is simply the expectation of the random variable Cost associated with a particular incident. *S* represents the set of all possible outcomes that might result from the occurrence of the incident. For each possible outcome *i*, the cost of *i*, namely the loss, is multiplied by the probability that *i* occurs. The summation of these products yields the expected value of the random variable Cost, which is the expected cost of the incident.

As defined earlier, exposure is the number of chances for an incident to occur. If a particular system has a chance of contributing to an incident each time it is operated, then the exposure from the system is the number of times the system is operated multiplied by the number of such systems in existence. When the system in question is used widely and frequently, this number can become quite large. For example, consider the in-flight breakup of TWA flight 800 over the Atlantic Ocean in 1996. The NTSB concluded that the probable cause of the accident was an explosion of the aircraft's center wing fuel tank, and the Board identified design issues affecting all Boeing 747 airplanes [14]. Exposure in this case would be the number of Boeing 747s in operation multiplied by the number of flights each aircraft would be expected to make in its lifetime. Given the popularity of the 747 and the near impossibility of surviving a commercial aircraft breakup at cruise altitude, the exposure and *E*[Cost] terms of the Total Risk equation would be very large, stressing the importance of implementing the Board's recommendations and reducing *P*[Incident Recurrence] in order to reduce the risk to an acceptable level.

The terms *P*[Incident Occurrence], exposure, and *E*[Cost] follow one's intuition in prioritizing incidents. Clearly, an incident with a high probability of recurrence with high expected costs warrants significant investigation, particularly if numerous systems are already deployed that might also be susceptible to the incident. Likewise, an incident with a small probability of recurrence, a low expected cost, or for which there are only a handful of susceptible systems that are rarely used might warrant only a minor investigation. Thus, Total Risk can be used as a metric to prioritize incident investigations, to determine where investigative resources would be best spent, and to decide which areas regulators, aircraft operators, and equipment manufacturers should focus on first when following up on investigators' recommendations.

As a second example of the use of Total Risk, consider the incident involving British Airways flight 027. It is very difficult to estimate the probability of recurrence but not impossible. The rates of failure of the relevant hardware components are probably known as is the rate of undetected damage occurring during maintenance. The cost of such an incident were it to result in an accident would be very high since there would be considerable loss of life and equipment. Exposure is also likely to be very high because of the prevalent use of TCAS. Thus, a rough estimate of the total risk could be calculated quickly and used as an indicator of the significance of the incident.

Follow-up Actions: A second important use of the concept of Total Risk is to guide the actions taken following an investigation. If Total Risk is high, then the follow-up actions should have a high probability of reducing it to an acceptable level. Many options are available to investigative and regulatory agencies and they need to be used carefully. At one extreme is the option of grounding the fleet and at the other there is the option of no action. In between, there are a variety of possibilities including required inspections, required equipment replacement, required equipment redesign, and so on. There are also options about how quickly any action should occur. Selection among options is a difficult activity if there is no effective mechanism for rating the seriousness of an incident.

Using British Airways flight 027 as an example once more, the actions taken following the incident were insufficient and fragmented despite the fact that Total Risk by the estimation above

was very high. Upon concluding its investigation, the CAA issued an airworthiness directive requiring air carriers using similar equipment to check and periodically inspect the equipment to ensure that it is functioning properly and notified other aviation regulatory agencies as well as equipment manufacturers of the problems it found. It also issued a recommendation to aircraft operators urging them to consider using other encoding schemes for transmitting altitude data since that was part of the problem. The CAA's recommendations did not require mandatory changes and the probability that they would reduce total risk to an acceptable level was small. More importantly, the report by British Airways contains useful insights about the incident yet it has not been made public nor led to appropriate general recommendations.

Iterative Reclassification: As an incident investigation proceeds, new details will emerge that affect the risk of future recurrence. The terms comprising the Total Risk equation will change as the breadth of possible event sequences is narrowed, faults are identified, and remedies are enacted. Consequently, new Total Risk assessments will periodically need to be made, and an investigation's priority relative to others will rise and fall as it is reclassified. After developing an initial set of recommendations, investigators might find that the risk associated with an incident has been reduced to the extent that their efforts would be better spent investigating other incidents with higher Total Risk assessments. Moreover, each reassessment will presumably lower the error in the estimate. Relying only on the initial Total Risk estimate is insufficient because this estimate is based on preliminary information and probably will not have a high degree of confidence associated with it. Therefore, in addition to the Total Risk metric for classifying incidents, a process is necessary to reassess incidents periodically in order to improve the confidence associated with Total Risk estimates.

Until an incident has been categorized, the initial Total Risk assessment cannot be performed, and the investigation into the incident should be given a high priority. Once assessed, the incident can be investigated according to its relative priority among other incidents. Investigators might then choose to reassess the incident on a strictly periodic basis (i.e. monthly or quarterly) or in light of major revelations concerning the investigation that might affect Total Risk, such as when a significant piece of evidence is discovered, when a defect is revealed, when a public inquiry is concluded, when recommendations are issued, or when remedies are implemented. Each reassessment will narrow the confidence interval on Total Risk. If reassessing an incident causes its Total Risk to increase, the investigation should be intensified until the risk is mitigated; if Total Risk decreases, resources can be diverted to more urgent investigations. The investigation may be concluded when investigators are confident that Total Risk has fallen below a predetermined acceptable level, which may depend on the incident's categorization, the type of operation (commercial vs. general aviation, scheduled vs. unscheduled), the flight rules in effect, the type of aircraft, and possibly other factors.

The goal of investigating incidents is to learn lessons that help to prevent the incidents from recurring. Some incidents might be symptomatic of severe defects that could lead to future casualties if not corrected; others could be fairly straightforward and involve accepted risks. By employing the risk-based metric and process proposed above, investigators might be able to determine more accurately which incidents have greater potential for teaching important lessons. Doing so would enable them to allocate resources first to those investigations that would likely have the greatest impact on safety. As a result, investigative agencies could begin to shift from a reactionary role in which loss motivates change to a proactive one focused on risk reduction.

Initial Total Risk Estimates: The Total Risk analysis as we have described above cannot be applied at the outset of an investigation because the data needed to estimate the parameters of the Total Risk equation will not yet be available; however in order to direct the allocation of

resources during an investigation's initial stages, it would be useful to have an estimate of Total Risk, albeit a very crude one. Although investigators will initially know little about an incident, they will have certain information from which a preliminary Total Risk assessment might be developed. This information includes the incident aircraft's flight plan, the type of aircraft, the stage of flight at which the incident occurred, the approximate time and location of the incident, prevailing meteorological conditions, the aircraft's last communication with air traffic control, and possibly preliminary statements from witnesses. These factors might be assembled into an Initial Total Risk Table containing precomputed standard Total Risk estimates compiled from historic statistical data. For a given incident category and set of circumstances, the table would provide estimates of the probability of incident recurrence, exposure, and the expected cost of the incident. Investigators could select which aspects of Total Risk to read from the table and which to estimate directly based on presently available information.

As an example of how the various ideas we have presented might be used, consider the British Airways flight 027 incident described above. Upon learning of the incident and categorizing it as a loss of separation between heavy aircraft, investigators would consult the Initial Total Risk Table using the categorization and other factors mentioned in the previous paragraph to obtain the initial Total Risk estimate. As the investigation progressed, investigators would rely less on the table and transition to estimating the Total Risk parameters directly, improving the accuracy of the Total Risk estimate in accordance with the objectives of Iterative Reclassification.

Remaining Work: The notion of Total Risk is a starting point for a metric that will allow investigators to assess the importance of incidents more accurately and allocate investigative resources accordingly. By assessing incidents based on the risks of future losses from their recurrence rather than their immediate losses, investigators can be more proactive in detecting safety problems before they contribute to accidents involving casualties or damage to aircraft.

Much work remains to be done before this metric can be put into practice. Because incidents are rare occurrences, estimating their probabilities is difficult. A model of cost will be needed to assess the expected loss associated with an incident that takes into account fatalities, serious and minor injures, and damage to aircraft and other property. Moreover, the estimation techniques and reassessment process presented here are intended to serve as examples and are quite preliminary. Before they can be applied to any investigation, they must first be developed more fully and tested on sample incidents to determine their precision. Statistics concerning incident rates and casualties decomposed according to incident type must be computed in order to estimate the parameters comprising the Total Risk equation. While similar statistics already exist, it is unclear whether they are in a form suitable for this purpose. Perhaps most importantly, investigators will need to set acceptable risk levels and establish criteria for determining which level would apply to a given incident.

Once these challenges are overcome, the estimation and assessment procedures would need to be refined so that they could be employed in the field quickly. Total Risk assessment is an overhead exercise and should not significantly detract from investigators' tasks of analyzing incidents and developing recommendations. While high precision cannot be expected from early estimates, they must be accurate enough to provide a rough indication of the worth of investigating an incident. Likewise, later assessments should help guide investigators in determining which aspects of the investigation to pursue next or whether to table the investigation and turn their attention elsewhere.

Conclusions

Commercial aviation accidents are serious occurrences that demand public investigations in order to correct safety problems and prevent future losses. Incidents are also important, however, since they often present the same opportunities to identify new lessons without the losses associated with accidents. Current accident classification schemes used by investigative agencies to allocate resources to investigations place too great an emphasis on the immediate loss from an accident and undervalue the importance of incidents with no loss. Consequently, incidents suggesting the presence of serious safety problems in onboard and ground-based systems are often ignored or not investigated with sufficient rigor to uncover these problems, which if left uncorrected could contribute to future incidents with more tragic outcomes. This dilemma was illustrated by the large disparity in the investigations conducted into the Korean Air flight 801 and British Airways flight 027 incidents. The latter received a much less rigorous investigation even though both incidents carried a high risk of recurrence.

Precedent existed for both of the incidents described in this paper. The Korean Air flight 801 accident followed a similar incident in 1994 that also involved a mis-configured MSAW system in which a Transportes Aereos Ejecutivos, S.A. Learjet crashed on final approach to runway 1R at Dulles International Airport approximately 0.8 nm short of the runway. The British Airways flight 027 incident followed a similar incident that also involved TCAS processing incorrect altitude data that occurred between two aircraft in January 1998 over Hawaii [3]. These prior incidents indicated the presence of serious problems with the manner in which the affected systems were designed or maintained; however the investigations either failed to address these problems or the follow-up actions were insufficient to correct them. As a result, opportunity remained for similar incidents to recur, *and they did*.

To mitigate this problem, investigators should reconsider the practice of classifying incidents based on their losses, and instead classify them based on the risk of future losses. Adopting risk-based schemes will allow investigators to be more proactive and address safety problems before they contribute to accidents with extensive casualties. For risk-based classification schemes to be useful, techniques will have to be developed for investigators to quickly assess the risk level of incidents early in the investigative process so that they can allocate resources accordingly.

This work was funded in part by NASA Langley Research Center under grants numbered NAG-1-2290 and NAG-1-02103.

References

1. Carley, William M. “Wires Crossed: Flawed Safety Device In Jets Gets Blamed For a Near Catastrophe.” *Wall Street Journal.* 12 October 1999, eastern ed.: A1.
2. Federal Aviation Administration, “Aircraft Accident and Incident Notification, Investigation, and Reporting”, Order 8020.11B. 16 August 2000. Washington, D.C.
3. Greenwell, William S. and Knight, John C. “What Should Aviation Safety Incidents Teach Us?” Technical report. CS-2003-12. Department of Computer Science, University of Virginia. 20 March 2003.
4. National Transportation Safety Board. “Accidents, Fatalities, and Rates, 2002 Preliminary Statistics, U.S. Aviation.” <http://www.ntsb.gov/aviation/Table1.htm>
5. National Transportation Safety Board. “Accidents and Accident Rates by NTSB Classification, 1983 through 2002, for U.S. Air Carriers Operating Under 14 CFR 121.” <http://www.ntsb.gov/aviation/Table2.htm>

6. National Transportation Safety Board. *Aviation Investigation Manual: Major Team Investigations*.
7. National Transportation Safety Board. *Controlled Flight Into Terrain, Korean Air Flight 801, Boeing 747-300, HL7486, Nimitz Hill, Guam, August 6, 1997*. Aircraft Accident Report NTSB/AAR-00/01. Washington, DC.
8. National Transportation Safety Board. *Public Hearing in Connection With the Investigation of Aircraft Accident, Korean Air Flight 801, B-747-300, Agana, Guam, August 6, 1997*. 24 March 1998. Honolulu, Hawaii.
9. Transportation Safety Board of Canada. "Investigation Process." (18 September 2002). <http://www.tsb.gc.ca/en/investigation_process/what_we_do.asp>
10. U. K. Civil Aviation Authority. "Hazardous Loss of Separation Between Two Aircraft Over Chinese Airspace." Doc Ref KMH/Pap/059, issue 1. 28 October 1999. London, U. K.
11. "Notification and Reporting of Aircraft Accidents or Incidents and Overdue Aircraft, and Preservation of Aircraft Wreckage, Mail, Cargo, and Records." *Code of Federal Regulations*. 2002 ed. Title 49, Pt. 802, p. 1195.
12. Aviation Safety Reporting System. "Program Overview." <http://asrs.arc.nasa.gov/overview.htm>
13. Johnson, Chris. "The Limitations of Aviation Incident Reporting." <http://www.dcs.gla.ac.uk/~johnson/papers/reminders/>
14. National Transportation Safety Board. *In-flight Breakup Over The Atlantic Ocean, Trans World Airlines Flight 800, Boeing 747-131, N93119, Near East Moriches, New York, July 17, 1996*. Aircraft Accident Report NTSB/AAR-00/03. Washington, D.C.
15. Federal Aviation Administration. "Advisory Circular: Aviation Safety Action Program (ASAP)." Advisory Circular 120-66B. 15 November 2002. Washington, D.C.
16. Leveson, Nancy G. *Safeware: System Safety and Computers*. Reading: Addison-Wesley. 1995.

Biographies

William S. Greenwell, Department of Computer Science, University of Virginia; Charlottesville, Virginia, U.S.A.; telephone - +1.434.982.2298; fax - +1.434.982.2214; e-mail - greenwell@cs.virginia.edu

William Greenwell is a graduate student at the University of Virginia. His primary interests include software dependability, incident analysis, aviation, and Japanese animation.

John C. Knight, Department of Computer Science, University of Virginia; Charlottesville, Virginia, U.S.A.; telephone - +1.434.982.2216; fax - +1.434.982.2214; e-mail - knight@cs.virginia.edu

John Knight is a professor of Computer Science at the University of Virginia. His primary interest is software dependability for safety-critical applications.

Elisabeth A. Strunk, Department of Computer Science, University of Virginia; Charlottesville, Virginia, U.S.A.; telephone - +1.434.982.2292; fax - +1.434.982.2214; e-mail - strunk@cs.virginia.edu

Elisabeth Strunk is a graduate student at the University of Virginia. Her primary interests are software systems dependability and practicing aikido.

Use of Incident Data Collection from Various Sources For Industrial Safety Performance Assessments

Nir Keren, T. Michael O'Connor, and M. Sam Mannan[1];
Mary Kay O'Connor Process Safety Center, Chemical Engineering Department, Texas A&M University System, College Station, Texas 77843-3122, USA
Phone: (979) 862-3985, Fax: (979) 845-6446, mannan@tamu.edu

Keywords: incident database, safety performance assessment, data collection, incident trend analysis

Abstract

Many organizations collect data on industrial incidents. These organizations differ from each other in their interests, data collection procedures, definitions, and scope, and each of them is analyzing its data to achieve its goal and to accomplish its mission. However, there were no attempts to explore the potential hidden in integrating data sources. Extensive efforts are required in order to integrate information from different data sources as well as to identify the effects of the individual aspects of data collection procedures on the quality and completeness of the data. This paper describes a methodology for incident data collection from various sources, and the opportunity that exists in a combined data mart for industrial safety performance assessment and identification of trends. Additionally, such analysis can be used to determine the areas for major reduction of losses and reduction in the number of incidents.

Introduction

There is an increased interest in using data on accidents to improve safety in the last 20 years. At the late 80s, V. C. Marshal consolidated incident data from sixty or so years and harnessed it toward loss reduction, and loss prevention in his book "Major Chemical Hazards" [1]. Today the interest is bigger than ever, because of the development of information technologies that looks promising in their abilities to see what "unarmed human eye" cannot see. Major efforts are being invested toward collection of incident related data. The US Department of Health and Human Services, The Agency for Toxic Substances and Disease Registry (ATSDR) maintain hazardous substances emergency events surveillance (HSEES) and publishes annual and cumulative reports [2], and is only one among many other type of data collection projects that is maintained by the Center of disease Control (CDC). The Department of transportation repository consist of large number of transportation safety Related databases, and many reports are available on their website [3]. The last are only two from at least 15 sources of information of incident related data that have been analyzed and incorporated in assessments of industrial safety performance by the Mary Kay O'Connor Process Safety Center, at the Texas A&M University, College Station Texas (MKOPSC). However, the main challenge in using incident related data only begins when the data is available.

Marono et al. suggests to use the European Commission accident-reporting database MARS as a support for the definition of a safety performance indicator system [4]. McCray and Mannan are the first to look in several databases for opportunities for risk reduction and loss prevention [5]. Mannan with O'Connor and West established the basis for a continual effort to exhaust the potential that is hidden in accident databases in their paper in reference 6. Mannan et al. looked again into EPA RMP Info database in order to determine the most significant chemical releases

[1] Corresponding author

[7] as part of the efforts described above. Early at 2002 the MKOPSC established a report on the feasibility of using federal incident databases to measure and improve chemical safety [8], and another report on assessment of the chemical safety in the United state for 2001 [9].

This paper presents the methodology that is being used by MKOPSC, the challenges, difficulties, measure, and shortly discusses future research and development to improve this methodology and increase its quality and capacity.

Assessment that is based on a methodology of incident data collection from various sources is a thorough process that has to be done carefully and in several stages. The flow chart in figure 1 is a simplified description of the process. The primary focus of industrial safety performance assessment, which uses the methodology described herein, is to establish a baseline metrics for the universe under investigation with regard to safety. This requires identification of incident trends, distribution of number of incidents, number of injuries, property damage costs, releases of materials, hospitalizations, and evacuations. These should be analyzed and correlated across the causes of incidents, equipment involved, initiation events, location, and other indicators. Several of the sources that are available collect only part or a sample of the information. However, it is possible to estimate the total number of chemical/product related incidents by applying statistical tools on the data. Implementation of indicator-based industrial performance measurement systems helps to determine whether the efforts invested toward safety improvement lead to the desired results. Other benefits are the ability to determine the areas that will lead to major reduction of losses and reduction in the number of incidents.

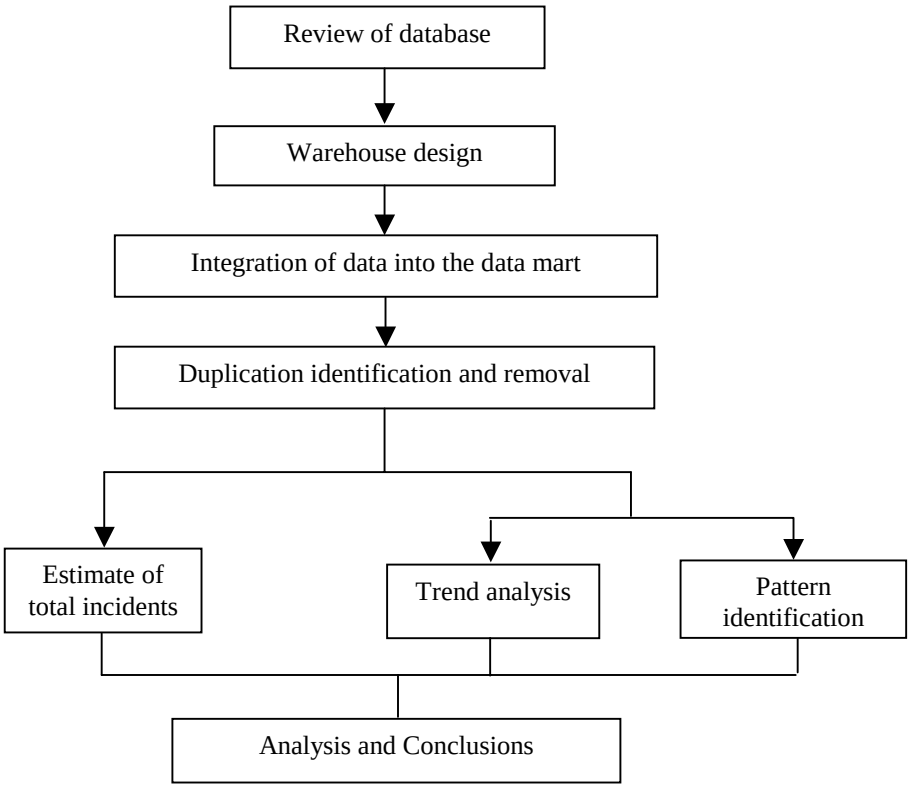


Figure 1- Methodology Flow Chart

Among the major conclusions from studies that have been conducted using this methodology is to not be "misled" by the amount of data that a certain source may consist of. In one study, a source of information provided about two-third of the data; however, it failed to collect significant data (e.g., failed to collect data with severe consequences). This conclusion justified the efforts that were required to broaden the search and combination of sources of information. A novel data collection methodology, based on News Clippings, has been established by the Mary Kay O'Connor Process Safety Center. This method uses search engines to query newspapers according to a predetermined set of keywords. The information is collected and submitted to the datamart. This method has several advantages including the ability to further investigate the incident or to verify the information if required.

Sources of Information

The process of integration of data from several sources requires a thorough analysis of the databases that collect information on industrial incidents. Table 1 consists of a list of more than a dozen databases from ten sources that were integrated for an assessment project for a certain industry sector. These databases were selected because they contain information that could be used to establish safety performance metrics for the industry sector.

The form of the data reflects the interest, purpose, and scope of the organization collecting the data. The lack of national and international standard of reporting incidents as Johnson mention in reference [11] has led to a lack of consistency among the sources with regard to coding used in the variety of fields. As a result, major efforts are needed to create an infrastructure that will allow data from variety of sources to "sit" together in a datamart. Figure 3 demonstrates the information flow until it reaches its final destination. At almost every node the data is being converted, and the process must be done diligently in order to avoid misinterpretation of the data.

It is important to emphasize that the sources do not release the information as it becomes available. A real-time data collection from various sources is a long process that takes at least three years, as can be seen in Figure 2. Because of its real-time nature, the news clipping data collection system creates several opportunities:

Table 1 - Sources of Information and Databases

Source	Database
Federal Emergency Management Agency (FEMA)	National Fire Information Reporting System (NFIRS)
U.S. Consumer Product Safety Commission (CPSC)	* National Electronic Injury Surveillance System (NEISS) * Death Certificates * Investigation Summary * Incident Summary
Mary Kay O'Connor Process Safety Center (MKOPSC)	News Clipping Database
States Associations	State of Iowa State of Florida
State Agencies	State of Texas
National Response Center (NRC)	Incident Reporting Information System (IRIS)
US Department of Health and Human Services, Agency for Toxic Substances	Hazardous Substances Emergency Events Surveillance (HSEES)

and Disease Registry	
U.S. Department of Transportation (DOT)	* Hazardous Material Incident Reporting System (HMIRS) * Integrated Pipeline Information System (IPIS) also known as Hazardous Liquid Accident Data (HLAD).
U.S. Environmental Protection Agency (EPA)	*Risk Management Program (RMP) 5-year Accident History *Accidental Release Information Program (ARIP)
U.S. Department of Labor, Occupational Safety and Health Administration (OSHA)	Accident Investigation System and several other databases.

- Development of procedures for incident investigation for the real-time data collection
- Identification of need for incident investigation and performing investigation
- Follow-up on information to validate causes of incidents and long-term consequences

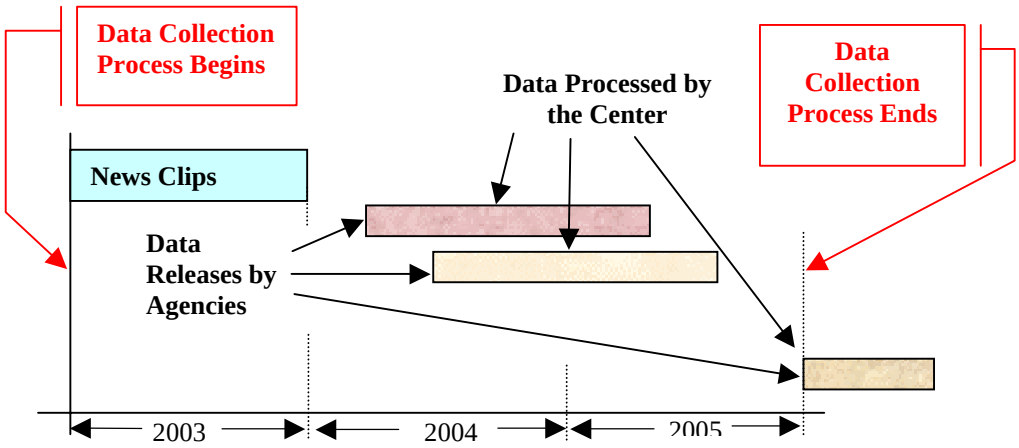


Figure 2 - Timetable of Real-Time Data Collection and Analysis

In many of the records that were examined, question arose that could have been answered quite easily if a real-time data collection process had been in place. In several of the records, it was hard to determine what is the cause, or what was the initiating event. As an example, one of the records contained data for an incident in Alaska.

The record indicated 99 fatalities for the incident. Since it is reasonable to assume that an incident with such large number of fatalities would be covered by the media as well as by incident investigation reports, a thorough research was conducted, which revealed that the incident actually resulted in a single fatality and 99 injuries.

Method of Duplication Identification and Removal

At the end of the data submission stage, it is required to identify duplications and to remove them. Johnson discusses many of the problems involved in automation of the process of duplication identification in reference [11].

There are two categories of duplications that are encountered during the consolidation of incident

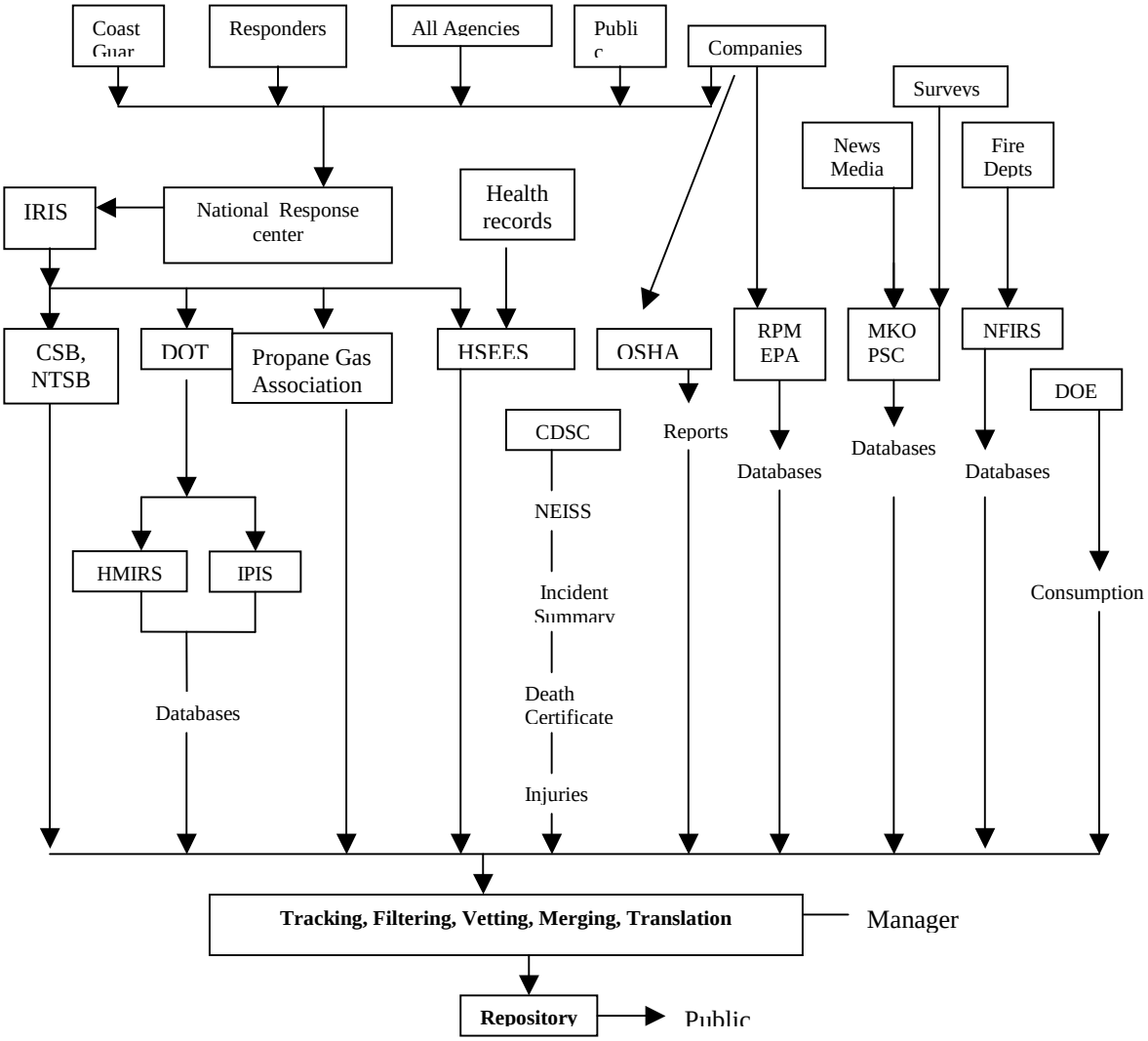


Figure 3 - Information Flow Chart

of incident information from a variety of sources:

- Duplications within the sources themselves, and
- Duplications among different sources

In general, it is much easier to identify duplications within the sources as compared to identifying duplications amongst different sources. However, the process of identification of duplications is similar in both cases. Duplication within the same source has the same type of information and is much easier to identify. The duplication identification process is illustrated in figure 4. The number of records in the list of 'Suspected as Duplications' is sensitive to the time frame that is employed. However, in order to verify that the time frame used is not arbitrary, the Center studied the sensitivity of the number of suspected as Duplications to the time frame.

As Figure 5 reveals, the number of incidents that are suspected as duplications is highly correlated with the width of the time frame (root mean square value of more then 0.98).

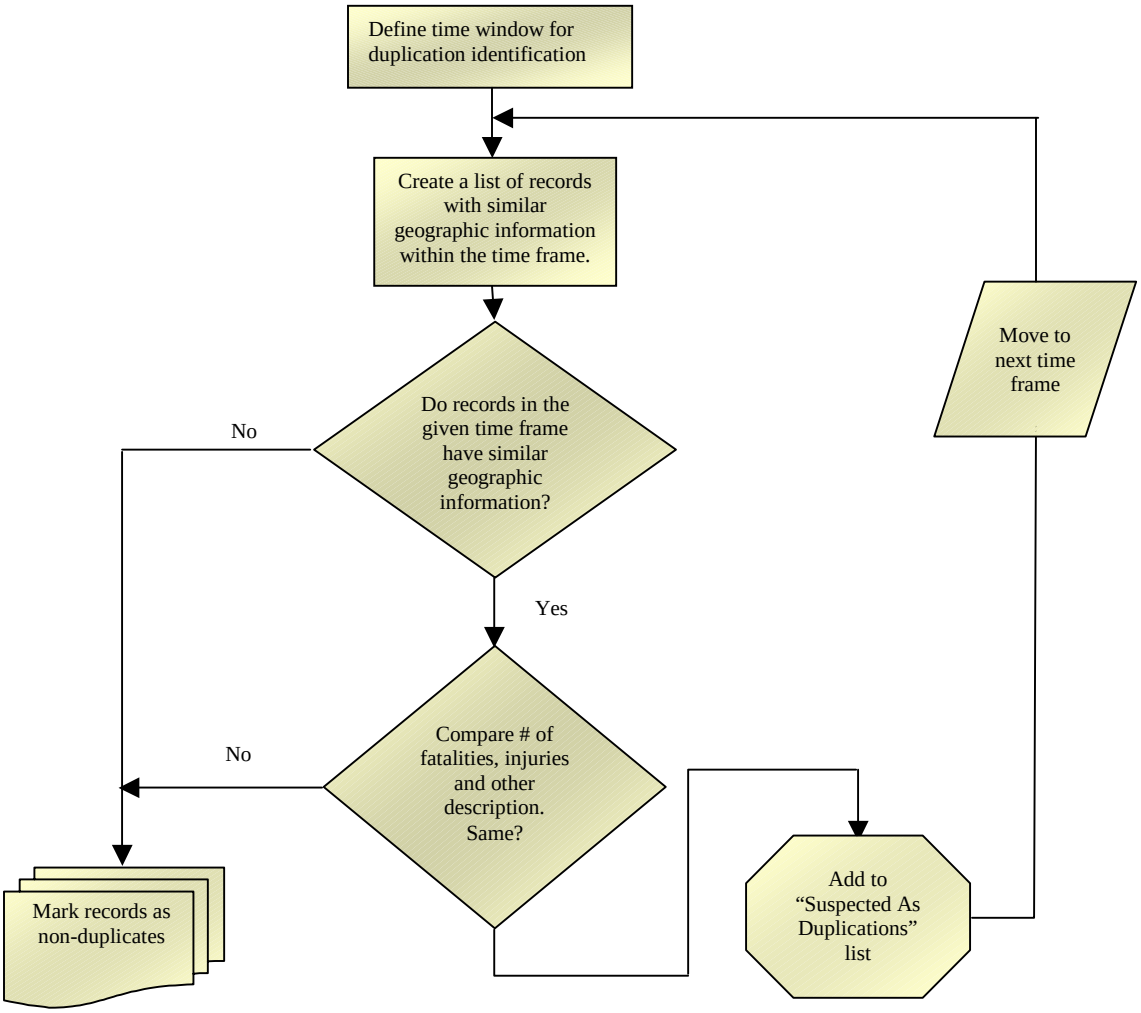


Figure 4 - Procedure for Identification of Duplications

The slope of the correlated line can serve as a qualitative relative indicator for the comprehensiveness of the database. Under the estimation that the probability of an incident to occur is not time dependent, the number of suspected duplication in a given time frame would increase as the portion of the universe of incidents increases. The slope of the curve becomes steeper as the comprehensiveness of the database increases.

Once the system creates a list of records that are suspected as duplications, the records are reviewed manually, and a decision with regard to these records are made. Records that are identified as duplications are marked, so queries will reveal only one of them. Identification of duplicates becomes quite difficult in cases where time of incident is not given.

As for duplicate identification within the databases, the process of verification of whether incidents are duplications varies according to characteristics of the incidents. NFIRS for example contains two types of duplications:

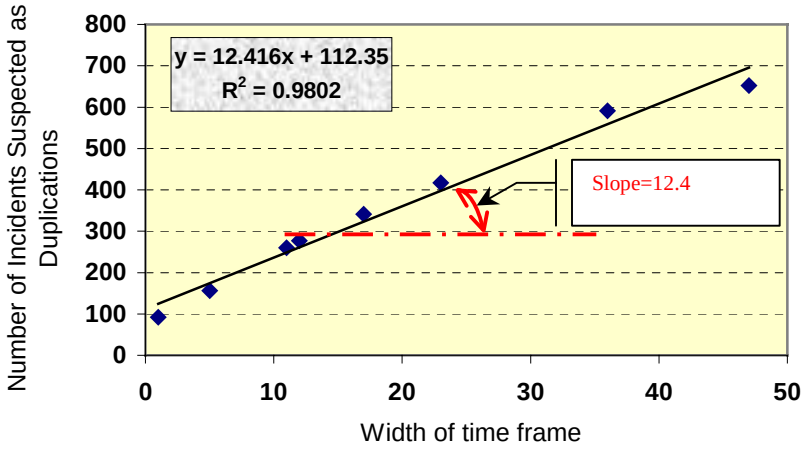


Figure 5 - Sensitivity to Time Frame Study

1) Fire department that reported the same incident more than once.
2) Incidents that were suspected as duplicates, because more than a single fire department reported the incident.

In the first case, the verification process is not complicated. In the second case, however, it was necessary to conduct an Internet search for county maps in order to determine if it is reasonable that a fire department from an adjacent county would assist another fire department and also report to NFIRS. In the majority of the cases the distance between the counties was too far to assume that the reports are duplicates.

An important criterion for identifying duplications is the number of injuries and fatalities. If two incidents that have other similar characteristics also show exactly the same number of fatalities and injuries, there is a high likelihood that one of these incidents is a duplicate. The system ignored incidents that have different number of injuries or fatalities. The Center applied manual checks and quality control procedures to ensure that duplicates were identified accurately and that non-duplicates were not eliminated inadvertently.

As for duplications amongst different databases, the process required relatively more extensive efforts, and each of the cases needed to be treated separately.

Methodology for Estimation of Total Number of Incidents

The process for estimating the total number of Industrial incidents in the Unites States can be explained by the theory of sets. Figure 6 illustrates the current situation. The gray area represents the total number of Industrial related incidents in the US. The white areas represent the actual number of incidents in each of the respective databases.

The number of incidents from each of the databases is a subset of the total number of incidents that this database would consist of if all incidents were reported to the source (the set). For example, NFIRS, which is a database that consists of reports from emergency departments, contains records from about 14,000 fire departments from 42 states. The records in NFIRS are a

subset of a set, which is the number of records that NFIRS would consist of if all 29,000 fire departments as well 6,900 emergency departments from the 50 states reported every Industrial incident to NFIRS. Figure 7 is an illustration of the relation between a set and a subset.

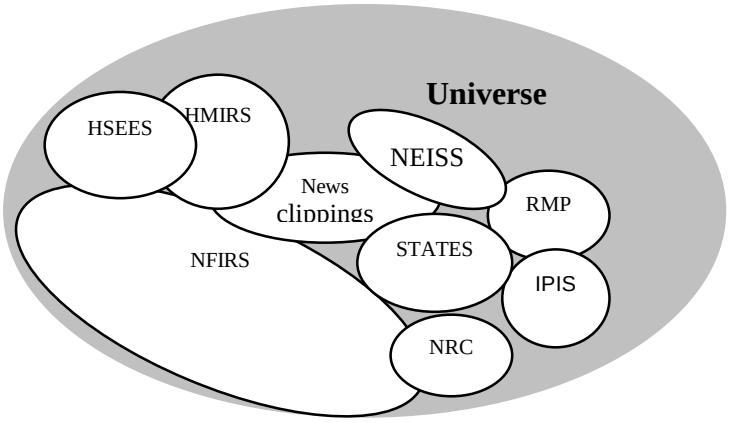


Figure 6 - Illustration of relation between total number of Incidents and Number of Incidents in the Sources

The Universe is a collection of all incidents that have the potential to be reported. Therefore, Universe is a composition of sets. The translation of the above to the theory of set language is as follows:

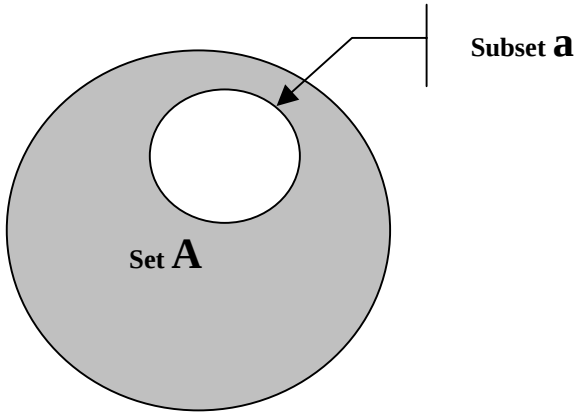


Figure 7 - Relation Between a Set and a Subset

$\mathbf{a_1}$ - is current records in database **DB1**
$\mathbf{A_1}$ - is the potential number of record in the database DB1, if all incidents targeted by this database were reported.

$\mathbf{a_1}$ is a subset of $\mathbf{A_1}$ ➔ $\mathbf{a_1 \subset A_1}$

$\mathbf{a_2}$ - is current records in database **DB2**
$\mathbf{A_2}$ - is the potential number of record in the database **DB2**, if all incidents targeted by this database were reported.

$\mathbf{a_2}$ is a subset of $\mathbf{A_2}$ ➔ $\mathbf{a_2 \subset A_2}$

The same principles applies to a_3, a_4,....., a_n or all the databases.

The Universe **S** is a composition of all the sets. However, there are overlaps among the sets, and therefore U is a union of the sets, as equation 1 shows:

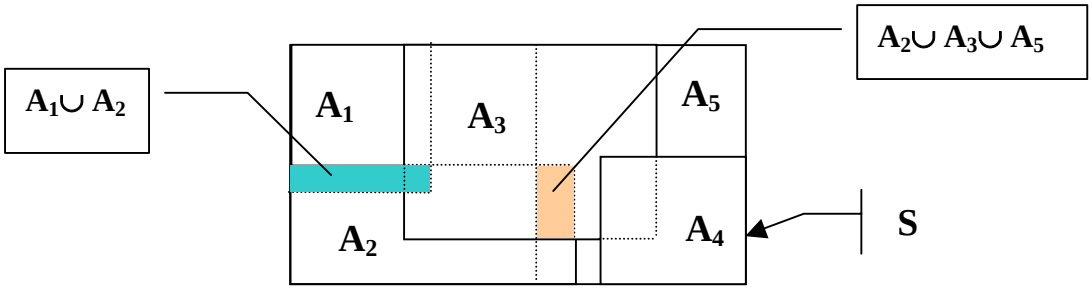


Figure 8 - The Universe is a Union of the Sets (Venn Diagram)

The sum of incidents from all databases prior to applying duplication identification procedure

The sum of the number of duplicates between every combination of pairs of databases

$$S = \left(\bigcup_{i=1}^{n} A_i\right) = A_1 \bigcup A_2 \bigcup A_3 \cdots \bigcup A_n = \sum_{i=1}^{n} A_i - \sum_{i=1}^{n}\sum_{j>i}^{n-1} \left(A_i \bigcap A_j\right) - \quad (1)$$

$$- \sum_{i=1}^{n-1}\sum_{j>i}^{n-2}\sum_{k>j}^{n-3} \left(A_i \bigcap A_j \bigcap A_k\right) - \ldots - \left(A_1 \bigcap A_2 \bigcap \cdots \bigcap A_n\right)$$

The sum of the number of multiplications among every combination of three databases

The number of multiplication that appeared in all of the databases

No multiplications found between more than two sources. Therefore, only the first two parts of equation 1 will be employed for the estimation purposes. These two parts are extended and are presented in equation 2:

$$\mathbf{S = A_1 + A_2 + A_3 + \ldots + A_n - [(A_1 \cap A_2 + A_1 \cap A_3 + \ldots + A_1 \cap A_n) + + (A_2 \cap A_3 + A_2 \cap A_4 + \ldots + A_2 \cap A_n)] + \ldots + (A_{(n-1)} \cap A_n)]} \quad (2)$$

The sequence of estimating the universe S is now simplified. The information that is available currently is the subsets a_i and the intersection between these subsets. Figure 9 presents the sequence of obtaining the information required to solve equation 2.

Following figure 9 is a description of the process for extrapolating the sets $\mathbf{A_i}$ according to the characteristics of each of the sources. The assumptions that were required in order to extrapolate the intersections between the sets are presented later.

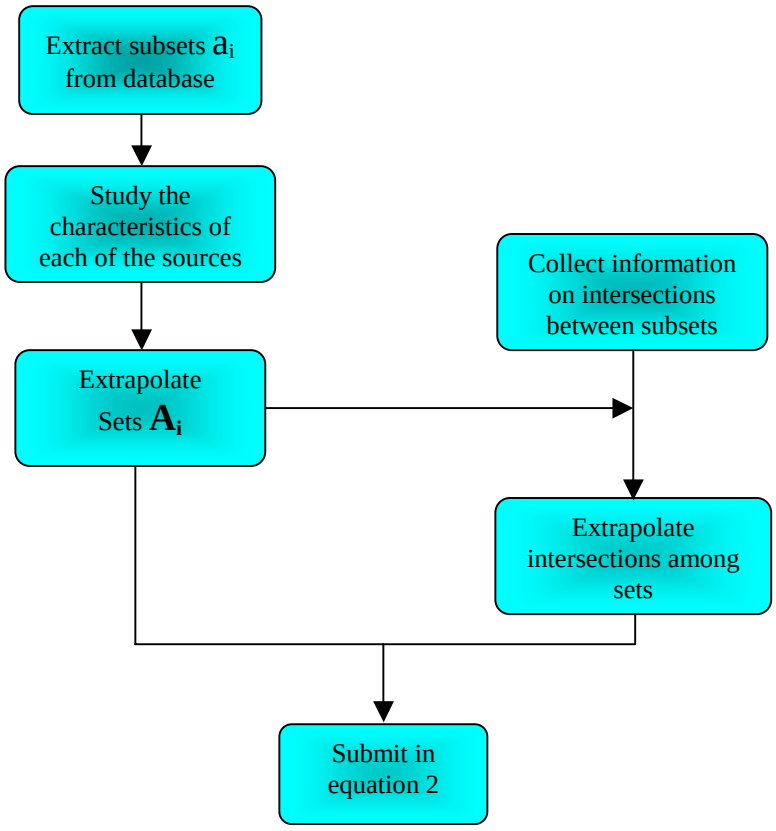


Figure 9 - Sequence of Estimation of Universe S

Extrapolation of Sets A_i: The purpose of collection of information is not the same for all the sources, and therefore the characteristics of each of these sources should be incorporated in order to calculate the number of incidents that the source database would consist of if it were to capture all the incidents that belong in its category. A set of considerations, as well as the methods for extrapolating of information from the sets, A_i, is developed individually.

Extrapolation of Duplicates: The ideal way to extrapolate the number of duplications is to sample several sample sizes of sub-sets and to identify number of duplicates for combination of sizes. By using this methodology it is possible to study how the number of duplications increases with increase of the size of subsets. However, in cases where the databases consist of relatively low number of duplicates, an approximation can be done by multiplication of the number of duplicates between sources by the ratio of the sum of the extrapolated number of the incidents and the sum of the actual number of incidents in the database.

Calibration of Information from Sources

When reviewing sources it is important to verify that the data represents the relevant population uniformly. The Consumer Product Safety Commission (CPSC) - National Electronic Incidents Surveillance System (NEISS) is a collection of injury data that are gathered from the emergency departments of 100 hospitals selected as a statistical sample of all 5,300 U.S. hospitals with emergency departments. NEISS surveys sample of hospitals that represent all ethnic groups and concentrations of population, and it is statistically valid to extrapolate this data. However, the Federal Emergency Management Agency, National Fire Information Reporting System (NFIRS) database, is a collection of incident reports from fire departments. About 30% of the fire departments in the US, from 42 states are report incidents to NFIRS. These fire departments vary in size range from departments that protect several dozen individuals (rural areas) up to departments that protect millions. In large fire departments all employees are paid, and in small fire departments employees are volunteers. The probability that large fire departments will report to NFIRS is much higher than small fire departments. Therefore, the analyst using NFIRS as a source of data should be aware of the distribution of consumption of the product that is under investigation. If fertilizers are the product in study, then the distribution of consumption in urban areas is expected to be much lower than the consumption in rural areas. However, rural fire departments in rural areas mainly employ volunteers and therefore the probability that these fire departments will report to NFIRS is low. In that case, information from NFIRS may be biased, and a calibration should be conducted. The Center used a survey of fire departments for calibration purposes.

Use of Indicators Toward Industrial Safety Performance Assessment

What are Indicators and What Do They Mean: As noted previously, a large amount of information exists about industrial incidents, including a large amount of information gathered by federal, state, and local agencies. The information gathered includes data on the specifics and numbers of releases of chemicals, on injuries, illnesses, deaths caused by chemicals and other products. Are any of these though accurate indicators of the state of effectiveness of chemical safety efforts? Do they tell us whether we are making progress in chemical safety?

An indicator is generally defined as an observed variable. Essentially, an indicator is presumed to reflect through a positive correlation a single underlying variable. The underlying variable being considered here is the safety of chemical processes. It is impossible to observe or measure industrial safety as a positive measure. It can only be measured as a negative measure, or an observable variable which is defined as when safety processes fail. The number of failures is an indicator, when taken in the context of the universe of potential failures, of industrial safety.

The indicator becomes more valuable in understanding the underlying variable when looked at over a period of time or as a trend. Trend analysis looks at an indicator or series of indicators over a period of time to observe if there is a general sustained movement of the time series upward, downward, or if there is no discernible pattern. Trend lines are used to graphically display trends in data and to analyze problems of prediction. Such analysis is also called regression analysis. By using regression analysis, it is possible to extend a trend line in a chart beyond the actual data to predict future values. The specific techniques that are most commonly applied include linear model, an exponential model, or a moving-averages model.

Trend analysis is commonly misapplied. For example, two or three data points cannot be used to develop a trend, though under a simple “eyeball” analysis it might seem so. In any trend and regression analysis, there always exists the assumption that a component of the underlying

variable is generated through a random or stochastic process interacting with the concrete set of data. Over a short period of time, the potential impact of this random process can be much larger than over a longer time period, where it becomes the "noise" or part of the error term in a regression analysis.

It is often better to use a number of different time periods in completing a trend analysis. For example, weekly measures viewed over a period of a year may indicate an upward movement of the number of injuries related to chemical releases. When viewed over a five-year period, the trend may be generally down, except for the current period, which could have been caused by an external variable such as a change in the definition of an injury, or a change in measuring techniques or methodologies.

On a larger perspective to be able to compare one set of indicators, for example for chlorine, to a set of indicators for petroleum products, the indicators must be normalized so that a comparison is made of essentially equal sets. Normalization is a general process by which two or more indicators are divided by an equivalent denominator. For the above example, an equivalent denominator might be the amount of chemicals produced. It is unadvisable to attempt to make a comparison across indicators that have not been normalized, as there is no actual basis for comparison and the resulting analysis is methodologically indefensible.

The Effect of Policies on Safety in the Chemical Industry: It is as importance to properly select the indicators, as it is to have an idea of what type of information you hope to see. The effects of changes in government regulations covering the chemical industry should be identifiable in the data.

If a specific policy change or new regulation has an effect on industrial safety, then graphic representations of the data recorded in databases would be reflected in the metric of interest. For example, the following graphic might illustrate the results of a governmental policy change. The performance in years one through five is relatively constant. During the fifth year (point A on the chart), a policy change is made and the resulting performance is shown by the value in year six (point B on the chart). It could be inferred the change resulted in about a 40 percent decrease in the number of incidents.

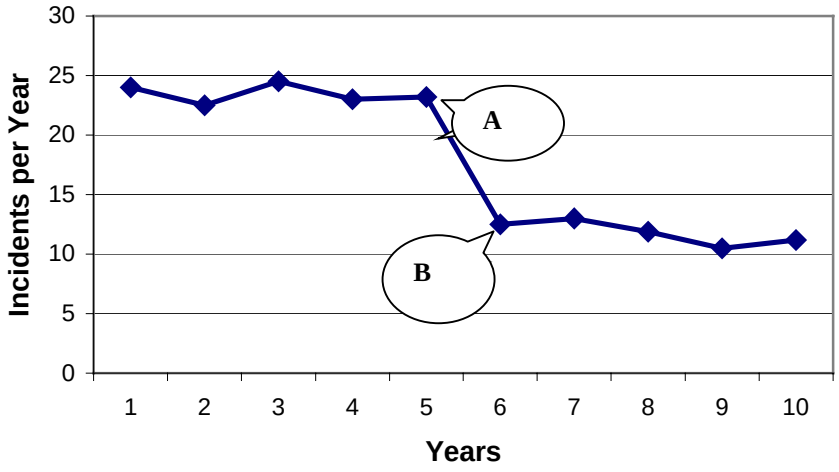


Figure 10 - Measuring the Impact of Policy Changes

Using indicators to investigate databases creates the opportunity to compare performance qualitatively among industry sectors. Figure 11 shows a plot of fatalities recorded by OSHA where a chemical is the primary or secondary cause of the fatality. Figure 12 presents a plot of fatalities from transportation related chemical incident. While figure 11 demonstrates a downward trend in fatalities resulting from chemicals, fatalities from transportation-related chemical incidents shows a slight upward trend over the 10-year period.

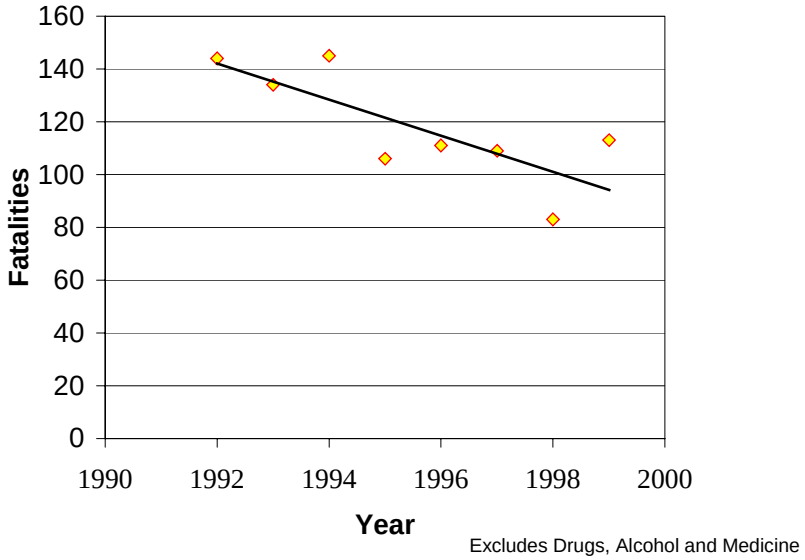


Figure 11 - OSHA Fatalities with Chemicals as Primary or Secondary Source

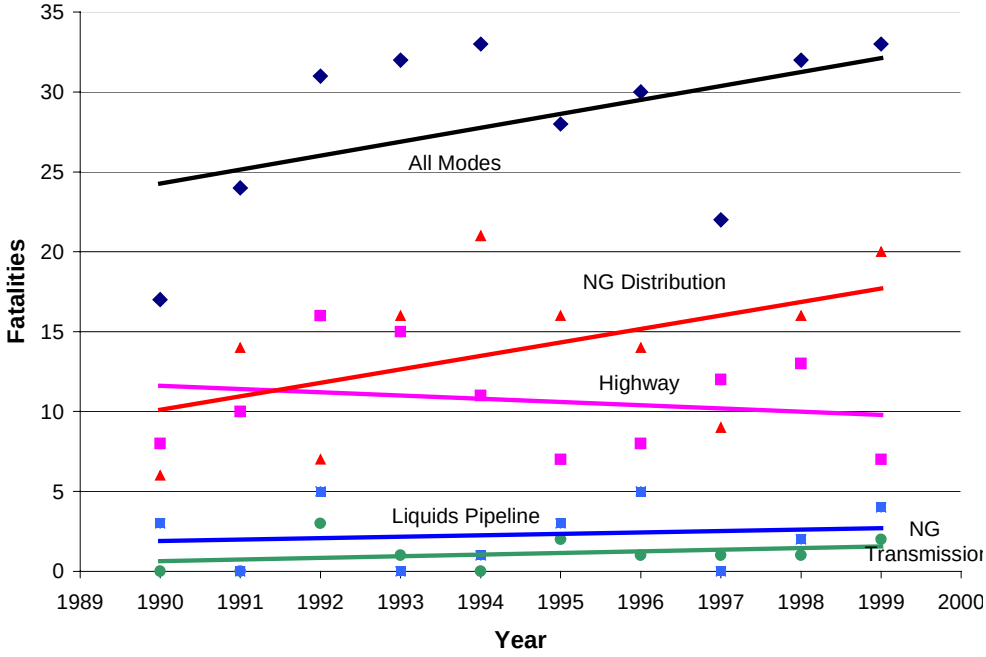


Figure 12 - Fatalities from Transportation-Related Incidents, Due to Chemicals

Figure 13 shows the result of implementation of data collection from various sources methodology on a petrochemical product for a certain year, and figure 14 demonstrates patterns of causes of incidents for the same product.

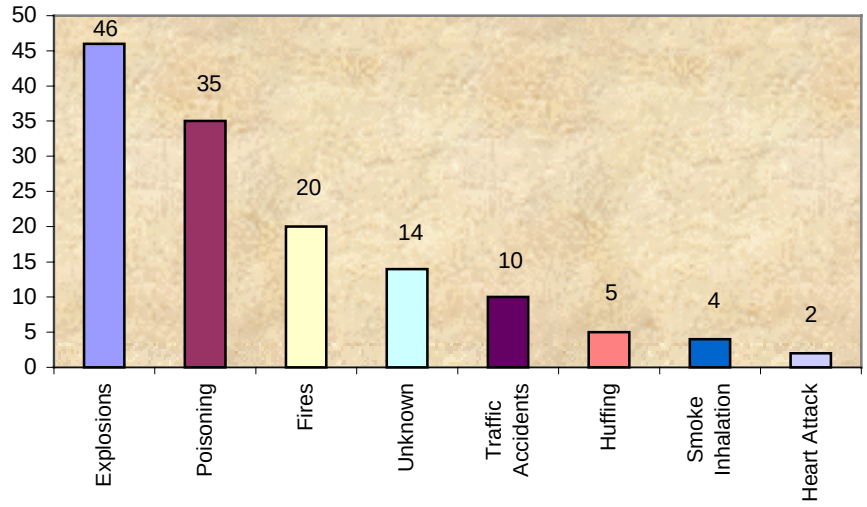


Figure 13 - Distribution of Fatalities by Cause of Death

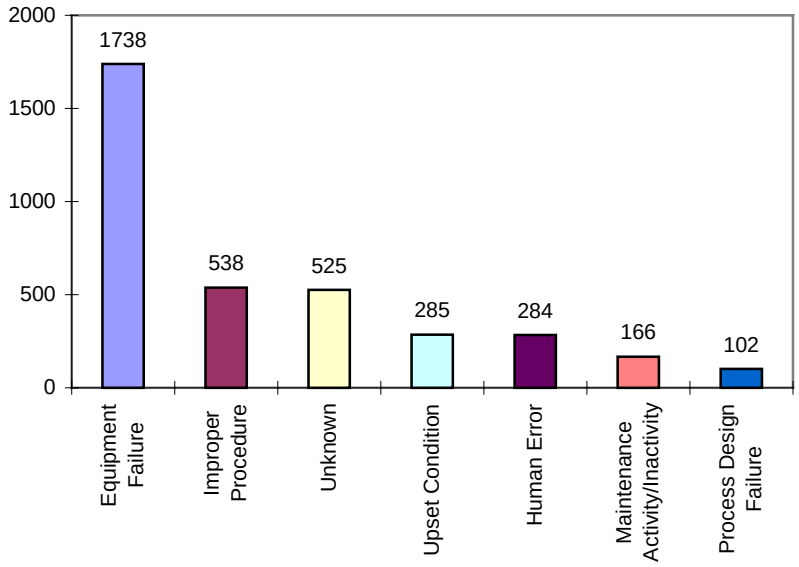


Figure 14 - Distribution of Number of Incidents by Causes

In general, collection of information of product is done in phases. First phase is implementation of the methodology on a single year in order to be able to learn difficulties that are related to collection of data on the product that is under study. After learning and implementing improvements that are required, a second phase is launched where data is collected for several years.

Further R&D

Extensive efforts are required in order to integrate information from the data sources as well as to identify the effects of the individual aspects of data collection procedures on the quality and completeness of the data. A holistic approach that suggests an innovation of tools, methods, techniques, and procedures in order to extract information from the large variety of sources of information is required. The principles as well as the methods of such a holistic approach will be applicable to many other disciplines such as civil engineering and insurance entities. The following is a short summary of the activities that will be conducted in order to automate the methodology that is described in this document: (1) development of methods for the detection and repair of common problems based on the individual storage formats; (2) development of techniques to allow identification and intervention of non-standard problems of these storages; (3) development of methods to automatically identify relational steps based on experts' seed knowledge; (4) merging the techniques above with current cleaning techniques to produce files that could be further processed without concern for storage of format irregularities; (5) development of methods for automatic integration of fields and relation building; (6) conversion of textual information into a schema of warehouse; (7) because of the multi-source legacy data, the development of duplication detection techniques as well as duplication handling techniques will be required as well; (8) allowing generation of flexible user application that prevents the need for involving content expert again, and yet ensure that subject relevance was maintained in the application output.

Conclusions

Several industrial safety performance assessments studies had been done by incident data collection and consolidation from a variety of sources. These studies demonstrated that it is worthwhile to collect data from variety of sources, and that much can be learned from the consolidated database. However, in order to accomplish the ultimate goals of safety performance assessment the consolidated database must include root cause information. In order to do that, a real-time incident data collection procedure must be established. Two major reasons for a real-time data collection process are: (1) news archives make the data available for a short period of time only; (2) the ability to further investigate an incident and get reliable results, decreases significantly with time.

There is enormous potential in employing data collection from a variety of information sources. This technique not only increases the amount of data captured by individual sources but also the ability to capture more diverse and significant incidents. In one of the studies, the methodology used by the Center resulted in the identification of 35% more incidents than the number of the incidents in the single largest source. In the same study, the methodology captured about ten times more fatalities than the single largest source.

The Center applied news clipping data collection procedure as a data collection methodology. However, this methodology is maximized only when applied in real-time because the data sources are available for limited periods of time, and because the Center can solicit additional useful information only during the period shortly after the incidents occur.

Analysis of the data identified a relatively low number of duplications. The majority of the duplications were found within the sources and not between them, i.e., the duplications are mainly because operators reported some incidents twice (or more). We believe that significant improvements can be made by real-time data collection.

References

1. V. C. Marshall, *Major Chemical Hazard*, Halsted Press: a division of John Wiley & Sons, NY, USA, 1987.
2. Division of Health Studies, Epidemiology and Surveillance Branch, Hazardous Substances Emergency Events Surveillance (HSEES): annual reports, http://www.atsdr.cdc.gov/HS/HSEES/
3. Department of Transportation – The Bureau of Transportation Statistics website: http://www.transtats.bts.gov/
4. M. Marono, R. Sola, J. A., J. Santamaria, *Using MARS database as Support for the Definition of a Safety Performance Indicator System*, Proceeding of the 10th International Symposium of Loss Prevention and Safety Promotion in the Process Industries, Stockholm, Sweden, June 2000, pp. 265-272.
5. Eboni McCray, M. Sam Mannan, *Use of Accident Database for Systematic Evaluation of Chemical Accident*, Proceeding of the Mary Kay O'Connor Process Safety Center annual Symposium, College Station, Texas, October 1999, pp. 47-161.
6. M. Sam Mannan, T. Michael O'Connor, Harry H. West, *Accident History Database: An Opportunity"* Environmental Progress (Vol. 18, No. 1), Spring 1999.
7. M. Sam Mannan, William J. Rogers, Fahad Al-Quarashi, *Analysis of the EPA RMP Info Database*, Proceeding of the Mary Kay O'Connor Process Safety Center annual Symposium, College Station, Texas, October 2000, pp. 39-45.
8. The Mary Kay O'Connor Process Safety Center, Texas A&M University, *Feasibility of Using Federal Incident Databases to Measure and Improve Chemical Safety, April 2002.*
9. The Mary Kay O'Connor Process Safety Center, Texas A&M University, *2001 Assessment of Chemical Safety in the United State, April 2002.*
10. Johnson, C.W., ***Using Case-Based Reasoning to Support the Indexing and Retrieval of Incident Reports***, Proceeding of European Safety and Reliability Conference (ESREL 2000): Foresight and Precaution, Balkema, Rotterdam, the Netherlands, 1387-1394, 2000.
11. Johnson, C.W., ***Software Tools to Support Incident Reporting Safety-Critical Systems***, www.dcs.gla.ac.uk/~johnson/papers/Safety_Science/software.html, 2001.

Biography

Nir Keren, Mary Kay O'Connor Process Safety Center, Chemical Engineering Department, The Texas A&M University System - 3574 TAMU; College Station, Texas 77843-3574, USA; Telephone - (979) 845-4950; Fax - (979) 458-1493; e-mail – mailto:nir@tamu.edu.

Nir Keren is pursuing a Ph.D. degree in the Interdisciplinary Engineering program at the Mary Kay O'Connor Process Safety Center, Texas A&M University, at College Station Texas. His primary interests include Industrial Safety Performance Measurements and novel applications of databases. Nir Keren completed his undergraduate studies at the Department of Mechanical Engineering at the Ben-Gurion University of the Negev, Israel. For 11 years, prior to joining the Center, Nir held positions as a Maintenance Manager, Project Manager, and Division Safety Engineer for several companies in Israel.

T. Michael O'Connor, Mary Kay O'Connor Process Safety Center, Chemical Engineering Department, The Texas A&M University System - 3122 TAMU; College Station, Texas 77843-3122, USA; Telephone - (979) 845-3489; Fax - (979) 845-6446; e-mail – mailto:tmo@che.tamu.edu.

Michael O'Connor is a Research Associate at the Mary Kay O'Connor Process Safety Center at

Texas A&M University. His research interests include high temperature heat exchangers and furnaces in ethylene and ammonia plants and metallurgy associated with these applications. Other areas of interest include incident database analyses and inherently safer design.

M. Sam Mannan, Mary Kay O'Connor Process Safety Center, Chemical Engineering Department, The Texas A&M University System - 3122 TAMU; College Station, Texas 77843-3122, USA; Telephone - (979) 862-3985; Fax - (979) 845-6446; e-mail – mailto:mannan@tamu.edu.

Dr. M. Sam Mannan is Professor of Chemical Engineering and Director of the Mary Kay O'Connor Process Safety Center at Texas A&M University. He is an internationally recognized expert on process safety and risk assessment. His research interests include hazard assessment and risk analysis, modeling of flammable and toxic gas cloud dispersion, inherently safer design, reactive chemicals and runaway reactions, aerosols, and abnormal situation management.

On Classification in the Study of Failure, and a Challenge to Classifiers

Kimberly S. Wasson; University of Virginia; Charlottesville, Virginia, U.S.A.

Keywords: classification, failure, investigation

Abstract

Classification schemes are abundant in the literature of failure. They serve a number of purposes, some more successfully than others. We examine several classification schemes constructed for various purposes relating to failure and its investigation, and discuss their values and limits. The analysis results in a continuum of uses for classification schemes, that suggests that the value of certain properties of these schemes is dependent on the goals a classification is designed to forward. The contrast in the value of different properties for different uses highlights a particular shortcoming: we argue that while humans are good at developing one kind of scheme: dynamic, flexible classifications used for exploratory purposes, we are not so good at developing another: static, rigid classifications used to trap and organize data for specific analytic goals. Our lack of strong foundation in developing valid instantiations of the latter impedes progress toward a number of investigative goals. This shortcoming and its consequences pose a challenge to researchers in the study of failure: to develop new methods for constructing and validating static classification schemes of demonstrable value in promoting the goals of investigations. We note current productive activity in this area, and outline foundations for more.

Introduction

The study of failure and the development and practice of investigation activities rely in part on a wealth of classification schemes. These schemes serve a number of goals and purposes, some of them more successfully than others. Common purposes include providing a springboard for consideration of ideas from many angles, through the filter of a classification scheme that facilitates such exploration, as well as providing a mechanism to group and organize low-level data for specific analytic purposes and to direct responsive action. These purposes suggest certain properties that allow classifications to be more successful at accomplishing their intended goals, and classifications that are useful for disparate purposes will embody disparate properties. For example, flexibility is desirable in some circumstances, while rigidity is desirable in others. Consistent interpretability is desired of all schemes.

In this paper, we first survey a number of classification schemes developed for various purposes relating to failure and its investigation, and abstract from this survey a continuum of goal types that such classifications are intended to promote. We then discuss classification properties that are useful or valuable in promoting these goals, as well as those that inhibit them. We argue that current practice is generally insufficient to achieve a particular set of goals. In particular, we find that humans are more successful at creating and productively using one type than another, and that our lack of strong foundation for development of the second type negatively impacts the value of data generated via the use of some schemes. We examine this issue to better understand its mechanics, and suggest how significant improvements can be made to the state of affairs. In particular, we encourage the systematic exploitation of relevant knowledge from related disciplines, and provide two models of how it might be done, in the form of examples of current productive activity in this area.

Survey of Classification Schemes

In this section, we examine several well-known and less-well-known classification schemes constructed for various purposes relating to failure and its investigation, and discuss their values and limits.

(1) Petroski's Design Paradigms: Overview: In [11], Petroski presents a collection of design paradigms that exemplify error and judgment in engineering. His goal is to highlight the role of judgment and experience in achieving good design, and through the presentation of case histories, he hopes to aid the development of judgment in his readers by providing them with years of experience essentially by proxy.

Value: Petroski provides an origin for a wealth of discussion, a scaffold for consideration of ideas from many sides, and a filter by which to draw out commonalities among many events (for example, by providing an example of "tunnel vision" in design, he encourages the reader to generate analogous additional examples, possibly from disparate subfields of engineering, in order to highlight cross-cutting concerns). From this, it is possible to gain insight by generalizing across large amounts of experience and extrapolate from patterns. The case studies also provide accessible cues for anecdotes that drive home messages. This is a non-trivial accomplishment--the lessons are sold and remembered.

Limits: Petroski's paradigms were never intended "...to constitute a unique, distinct, exhaustive, or definitive classification of design errors." Indeed, they can sometimes be almost too flexible, so as to have little meaning (if everything can be everything, what is anything?) Thus this classification is not appropriate for doing any sort of quantitative analysis, but neither is it meant to be. As for the goal of aiding in the development of judgment, the paradigms are light on mechanism. That is, we are provided with a collection of models of good and bad judgment, but it is never explicitly discussed just *what* judgment *is*, at a psychologically low level, and how this classification scheme helps to develop it.

(2) Perrow's Interaction/Coupling Chart: Overview: In [10], Perrow argues that there exists the "possibility of managing high-risk technologies better than we are now," in addition to obvious steps like safer design and better operator training [10]. He argues that even with the most advanced safety mechanisms in place, some kinds of accidents are inevitable. He characterizes systems susceptible to such accidents by their high interactive complexity, that is, a large number of dependencies among elements of the system, and their tight coupling, that is, a lack of flexibility in the structure and timing of the processes that make up a system. High interactive complexity and tight coupling together affect the behavior of systems possessing them in critical ways that make appropriate response exceedingly difficult in critical situations.

Value: Perrow provides a scaffolding for discussing salient characteristics of high-consequence systems. It is somewhat less flexible than Petroski's paradigms, partly because Perrow wants to be able to drive policy decisions based on his classification, and to do so, it must have some integrity. He intends to provide a foundation for decision-making about which kinds of proposed systems should and should not be built, and which kinds of existing systems should be abandoned or modified. His classification does inform such decisions with useful information not previously thus synthesized.

Limits: Perrow only succeeds to the point that one agrees with his rationales for assigning industries to classes. The classes *are* somewhat subjective. While interactive complexity and tight coupling are reasonably well-defined notions and definitely capable of generating insight about

systems, they do not necessarily provoke easy consensus in classifying systems under consideration when other issues are thrown in that can affect policy decisions. In particular, there is disagreement about some aspects of the safety of nuclear power (but such disagreements provoke discussion, an emergent value from this limit).

(3) Reason's Generic Error-Modeling System: Overview: In [13], Reason discusses the psychological characterization of human error and presents a classification scheme by which to organize human error types. It is based on Rasmussen's SRK model [12], and enhanced by Reason's addition of further distinctions. The Generic Error-Modeling System uses research results from psychology about the mechanics of human information processing to inform a breakdown of error types according to the cognitive processing mode with which they are associated. It takes as a substrate the notion of the mind as a General Problem Solver (as per [9]) and first separates error events according to whether they occur before a problem is detected or afterward. Those that occur before map to Rasmussen's Skill-Based level, while those that occur after map to his Rule- and Knowledge-Based Levels. Those occurring before are further divided into slips and lapses, and those that occur afterward separate into Rule-Based mistakes and Knowledge-Based mistakes.

Value: Reason provides an explicit examination of mechanics that is theoretically founded and can be used to motivate preventive and corrective actions. It is not only more objective than the schemes of Petroski or Perrow, but it is more likely to be meaningful to the creation of strategies explicitly intended to take this problem into account when designing systems that better cope with it.

Limits: While it provides the possibility for constructing useful responses, it doesn't actually follow through (though that is reasonably beyond the scope of Reason's work). It lacks functional direction for application and requires others to take up the charge. One such researcher is Busse, whose work will be treated later in this paper [1].

(4) NASA, FAA, AIMS and ESRD Classifications Schemes for Use in Investigation and Monitoring: Overview: We treat these classification schemes together because they have in common certain properties with which we are concerned.[1] The schemes under consideration here are drawn from NASA's Procedures and Guidelines for Mishap Reporting, Investigating, and Recordkeeping [6] (NPG), the FAA's Order on Aircraft Accident and Incident Notification, Investigation and Reporting [16] (FAAO), the Australian Incident Monitoring Scheme [14] (AIMS), and National End Stage Renal Disease Patient Safety Taxonomy [7] (ESRD). The NPG and FAAO each outline policies and procedures governing activities to be undertaken during the investigation of specific incidents and accidents under their respective jurisdictions; the FAAO additionally governs certain activities of the National Transportation Safety Board (NTSB). Included in it are a number of schemes that direct the classification of the large volume of information generated as a result of an investigation. For example, the NPG and FAAO each provide a scheme for classifying undesired events (NASA "mishaps", aircraft incidents and accidents) according to severity; these classification assignments drive organizational response. The AIMS and ESRD Initiative provide direction in monitoring of ongoing activities and events and the inclusion of relevant information in databases for analysis. For example, both systems include schemes for attributing various levels and sources of causality of an adverse event.

[1] The reader might recognize that this is itself an implicit classification. It has its own value of aiding in the organization of this work and drawing the reader's attention to properties common to the classifications under discussion, and its own limit of being ad hoc, that is, useful for the purpose at hand, but in focusing on particular properties, it potentially ignores others by which other useful analyses might be attained.

Value: All of the classification schemes provide guidance in accomplishing activities that have the potential to bring about improvements in the systems with which they are concerned, through correction and prevention of existing faults and other sources of failure. Thus they apply at a lower level than the schemes previously discussed, which lack specific application mechanisms. Further, the results of the classifications associated with individual investigations and monitoring activities can be collected and analyzed together for trends that can provide additional insight.

Limits: The main drawback to these schemes is that inherent in them are semantic ambiguities that impede many of the goals of investigation and monitoring. For example, if an ambiguity allows two different reporters to classify identical events in different categories, then the classification scheme lacks integrity: analyses based on it are likely to find false patterns and miss actual ones. For example, the NPG classification scheme for mishap severity allows an interpretation with a contradiction [4], the FAAO generally classifies events involving loss of life as accidents while classifying those that involve hazardous materials, even if loss of life occurs, as incidents [16], AIMS gives little guidance in teasing apart the vagaries of inattention, fatigue, haste, or stress [1, 14], and the ESRD taxonomy definitions of root cause, proximate cause, and proximal cause are so circular and ungrounded as to leave the user more confused than had he not read them [7]. These ambiguities exist because the classification schemes were not developed with, for example, the rigor of Reason in exploiting a scientific basis (but they do have clear applicability where Reason lacks it). Busse characterizes the AIMS classification as forcing reliance on judgment and lacking in substance or discriminatory power, which can be said of each of the other schemes as well. The FAAO indicates explicitly that in the case where a particular need is not provided for by the document, investigators should use their judgment. But ambiguous classification schemes and over-reliance on judgment cannot promote the goal of integrity in classification, and thus meaningfulness of analyses and validity of responses.

(5) TransportNSW and NTSB Classification Schemes for Use in Investigation and Monitoring:
Overview: We treat these classification schemes together because they have in common certain properties with which we are concerned. The schemes under consideration here are drawn from The New South Wales Department of Transport's monitoring system for Signals Passed at Danger [15] (NSW) and the National Transportation Safety Board's scheme for allocating investigative resources according to distinctions in event severity [2, 8] (NTSB). NSW distinguishes three levels of severity of signals passed at danger: low, medium, and high, using factors such as total distance by which the signal was overrun and whether damage or death resulted. NTSB separates major from serious accidents using similar factors such as amount of damage and number of fatalities.

Value: Each of these schemes is unambiguous, and therefore capable of providing consistency that is missing in the schemes treated in the previous section. Patterns and trends observed are more likely to be actual patterns and not false ones.

Limits: While the disambiguity of these schemes allows more consistent tracking of data, it is not clear that the data being tracked are interesting. This is because the divisions among the classes in the schemes are based on observed outcomes rather than the origins of those events. In order to respond in a useful way, it is necessary to know how an event came to pass and not just its result. Whether a train overruns a signal by 183 meters or 184 meters (two separate categories in the existing scheme) is far less useful than knowing, for example, the distance the train is likely to be carried by its mass and inertia once the brakes are applied. The latter could form the basis of a taxonomy that helps to distinguish whether the brakes were a factor in an undesired event. Likewise in separating aircraft accidents from incidents based on the severity of the loss

sustained; whether a loss was sustained is more often a function of chance or luck than of the origins of contributing faults. [2] describes a near miss that would almost certainly have resulted in a midair collision had the aircraft had GPS installed; as it was, that the two aircraft missed each other was attributed not to any safety measure but rather to random noise in the ability of the aircraft to follow their programmed flight paths [2]. Certainly in observing outcomes there are intuitive apparent differences: multiple deaths seem to warrant more scrutiny than minor mechanical damage, and factors such as public relations encourage this to be so. But this is a false correlation when it comes to strategizing for prevention: in each of these cases, the quantitative measure of a degree (of loss, of damage, of arbitrary distance overrun) masks the problem of determining, through qualitative means such contextualizing an overrun distance in something physically meaningful, the likelihood of recurrence ([2] for aircraft near miss incident).

Analysis

The survey presented above affords the description of a continuum of uses for classification schemes, that suggests that the value of certain properties of classifications is dependent on the goals the classification is designed to forward. For example, flexibility in a classification might be desirable if the scheme is intended to provide a springboard for exploration of ideas, as in directing the consideration of an entity from many sides (as with Petroski's paradigms). On the other hand, rigidity is more desirable if we are concerned with trapping data related to a particular event into a characterization to be analytically processed with the goal of producing specific, actionable results (as with the tracking of error data to be used in informing, for example, system redesign). This continuum can thus be partitioned to reflect a dichotomy whereby non-domain-specific, high-level classifications tend to be dynamic and flexible, based on intuition, and in the service of exploration and generation of insight, while low-level, domain-specific classifications, generally applied to specific events under investigation or monitoring, tend to be static and rigid, and in the service of creation of analyzable organization in data, in a repeatable fashion.[2] One classification type favors flexibility, the other favors consistency and integrity. Among the classification schemes treated in this survey, those presented in survey sections 1 and 2 are more representative of the first type; Petroski and Perrow are concerned with abstracting inductively from large numbers of events in order to intuit patterns worthy of exploration. They might encourage some analysis, but neither provides much in the way specific, low-level results to be acted on in the correction or prevention of domain-specific faults. The schemes presented in the final two sections of the survey are the complement; they are explicitly constructed to trap low-level details of specific events in order to collect and analyze them, to direct corrective action on the systems involved. The remaining scheme, that of Reason, is something of a straddler in this analysis; while his scheme is not domain-specific or in the service of investigation of individual events, it *is* concerned with low-level cognitive mechanics that precipitate human error, and in addressing these mechanics, provides the foundation for specific corrective strategies in systems that suffer as a result of human error. This attention to origination of faults (in this case, human errors), and not just observation of their results, is valuable and precisely the kind of insight lacking in the NTSB and NSW schemes that distinguish classes by more arbitrary or less meaningful factors. However, what Reason lacks is the domain-specificity and application

[2] Repeatable, because we desire consistency not just within the investigation of a single event, but across multiple investigation instances that can be analyzed together in studies of wider scope. Further, repeatability allows analysis of the process itself in order to improve it; one cannot improve on a process that one cannot characterize and document, to know where to start in making the improvements.

mechanism to be able to *use* this low-level, mechanical information to inform corrective strategies in specific systems.[3]

The contrast in the value of different properties of different classification schemes for different uses highlights a shortcoming in achieving a particular purpose in the continuum: we argue that while humans are good at developing one kind of scheme, they are not so good at developing the other. Being furious pattern matchers, we are good at spotting the things common among entities under consideration; this would seem to indicate that we might do well at all aspects of classification. However, while we are able to induce patterns in disparate entities, and do well in exploring ideas and generating insights through dynamic, flexible classifications, and work with (and benefit from) their ambiguities and contradictions, we are far less successful at reaching the goals we intend for rigid, static, domain-specific classifications. We can construct them, that is, propose *some* taxonomy for a given environment, and declare it to be rigid and static, but it often turns out to be the wrong set of divisions--invalid as a rigid system, because we failed to set it up along the best possible lines and with the necessary explicit precision available to users. Without these properties, such schemes cannot meet the goals of integrity, meaningful analyzability of data, repeatability, or ability to motivate valid corrective responses.

Specifically, these deficiencies derive from two sources. As we saw in survey section 4, achieving the necessary explicit precision is one problem. This is a linguistic issue, and derives from the fact that our needs for this kind of precision are not something we are cognitively built to handle naturally. In [4], Hanks, Knight, and Holloway discuss the specific cognitive mechanics that allow for ambiguity and thus provide the environment for assumption to be relied upon in interpreting language. However, while these mechanics provide for high-bandwidth and language efficiency in the common case in which interlocutors share sufficient experience, these same mechanics backfire with severe consequences when the needs for precision are out of the ordinary. This allows for, and more likely, encourages, variation in the interpretation of, for example, guidelines directing the investigation of any disaster within their scope, limiting the degree to which that investigation can achieve its goals.

Survey section 5 provided discussion of the other main problem with developing successful static classification schemes; even if we can achieve the requisite precision to allow all users to arrive at the same interpretation, such interpretations are only useful if they are meaningfully connected to determination of origins of faults and not just their results. Recall, it is of far more value, from a standpoint of correction and prevention, to know if the distance overrun by a train was a factor of the braking system than whether it was 183 meters or 184; likewise is of far more value to know that two aircraft events of the same *potential* severity derived from the same electrical malfunction than that one of the events was accompanied by actual damage and loss of life while the other was not.

Why aren't we good at building valid static classifications? Because we build them on the wrong bases and with insufficient rigor in disambiguation. Our lack of strong foundation in developing useful rationales and methods of accurately increasing precision impedes progress toward a number of goals, like repeatability of process, meaningful analysis, and ability to drive valid corrective action. The problem has occasionally been referred to as an issue of the integrity of the classification, but most existing solutions amount to little more than "be careful." "Be careful" isn't enough. We need foundations. We cannot escape all uncertainty in interpretation, nor can we know every rational path from origin(s) to fault, but to advance our ability to generate useful

[3] We recognize that this is quite reasonably beyond the scope of his work; it is rather the field that lacks the means to apply Reason's work. Busse is making strides in this direction [1].

responses to specific events, and thereby to advance our understanding of failure generally, we should be trying more systematically to use all resources at our disposal to direct ourselves in removing all *unnecessary* uncertainty and misguidance.

Mandate

This shortcoming and its consequences pose a challenge to researchers in the study of failure: to develop new, more rigorously grounded methods for constructing and validating domain-specific static classification schemes. It is not enough to "be careful" in writing precision-oriented guidance documents, nor is it sufficiently productive assign investigative resources or develop corrective actions based on the results of a fault without also accounting for its origin(s) and the potential damage they allow. Even if we attempt to account for these, we are not doing as much as we can unless we are applying available relevant results systematically. We need methods of rigorously analyzing domains to access the structure and organization that mediates the knowledge through which we actually interact with the domains. It may be that one intuitive User Interface failure mode of a device is having its power supply kicked out of the wall, while another is having a coffee spilled on it, but these scenarios do not tell the whole story, do not represent the whole picture of our interaction with this specific device and the organized collection of concepts and understandings that mediate this interaction. Can we do better at capturing this information and driving static classifications off of it, such that the classifications have more integrity and thus the data analyses generated from them are more meaningful and the processes themselves can be made rigorous and repeatable and corrective actions are valid?

Current Work in Advancing this Cause

There are at least two projects taking specifically this approach in developing more valid static classifications. One is the methodology for better management of natural language throughout system lifecycles advocated by Hanks and Knight [3], which provides not only for better organization and contextualization of domain terminology for use in investigation guidelines and report documents, but for virtually any other component of a system lifecycle that relies on the use of natural language. This methodology is founded on results from linguistics and cognitive psychology that characterize specific cognitive mechanics involved in communication, and uses these mechanics to inform well-defined techniques and support tools for reducing the potential for miscommunication embodied in lifecycle artifacts using natural language. In particular, it addresses the related problems of precision and accuracy in communication using domain-specific terminology, to be used in classifications or otherwise. That is, it provides support structures and direction for communicating the correct concept, and at the appropriate level of granularity. Cognitive linguistic research results are thus exploited to shape methods that can be used to drive the construction of classifications that are less ambiguous and more cooperative with the deficiencies of natural human semantic organization--these methods do not just add explication, they add it in the right amounts and in the right places to allow interpretations and therefore dependent decisions of higher integrity.

Another project exploiting existing foundational results from relevant areas is the Cognitive Error Analysis methodology of Busse [1]. This work seeks to use existing results in psychology and cognitive science to inform techniques designed to reduce the incidence of human error in critical environments. It starts from the recognition that Reason's classification scheme, as discussed above, has desirable rigor in examining the origins and mechanics of human error, but lacks sufficient direction in application of its insights to the problem of developing preventive and corrective strategies. Busse addresses the problem of providing that functional direction, and her work has led to examination and improvement of a number of classification schemes used in

critical environments (e.g., a neo-natal intensive care unit). In particular, her work "...shows how error categorisation, when done according to a cognitive level of performance and latent factors, can provide the basis for sound, structured, and theory-based remedial recommendations" [1].

Of note is a further project that appears promising. The Laboratory of Decision Making and Cognition at Columbia University has a project in Human Error in Naturalistic Medical Environments. Among its goals is: "to develop a cognitive framework of medical errors in critical care environments, where decisions are often made under high stress, time pressure, and with incomplete information, which leads to a high degree of uncertainty in diagnosis and management. Our objectives include (1) developing a cognitive taxonomy of errors where each type of medical error is associated with a specific cognitive mechanism (2) a theoretical explanation of why these errors occur and prediction of the circumstances in which a specific error could occur, and (3) a cognitive intervention strategy based on the taxonomy that can prevent or reduce each category of medical error" [5]. However, while this initiative appears well-founded with regard to the priorities discussed in this paper, there are as yet no apparent results from this research group.

These projects, while providing example models, on their own contribute only drops in the proverbial bucket; their value has not yet been exploited, and there is a wealth of other foundations that can be explored for their usefulness in constructing more valid and useful static classification schemes. For example, there is far more available in both psychology and linguistics than either Busse or Hanks and Knight have explored. Among further options in linguistics is discourse analysis, and there are any number of high- and low-level psychological results relevant to human information processing, problem solving, and memory with surely hidden value. Sociology can inform interactions among individuals in modes other than linguistic communication. Biology can inform meaningful classification schemes for analyzing the effects of devices on live tissue. Chemistry and physics can do the same for interaction among any bits of matter or energy. In theory, results in the natural and social sciences could obviate the need to rely on any ambiguous or ungrounded classification scheme, but we as a community must make the commitment of resources to apply them.

Conclusion

There exists a continuum of uses and goals for classification schemes in the study of failure, and thus a continuum of properties that are useful and desirable in these schemes. The continuum of properties can be partitioned into a dichotomy opposing schemes that are dynamic and flexible, used for exploration and discovery, vs. those that are static and rigid, used for trapping data for analysis and creation of specific new results that inform preventive and corrective actions. The first type characterizes domain-independent inquiry, abstractions from many events, collected according to observed patterns that encourage new consideration of new angles. Flexibility is valuable, since it allows the examination of entities from many sides, sometimes simultaneously, and encourages generation of insight. The second type characterizes domain-specific inquiry, and in-depth investigation of individual or closely related events, in which classifications are created and applied, rather than observed and induced. The goal of the second type of classification is the opportunity for meaningful analysis, repeatability of process, integrity of results, and the ability to act on them. While humans are successful in creating and using the first type of classification, we are not as good at meeting the goals of the second type, because even though we can make schemes rigid by fiat, we have difficulty in developing classification schemes that are sufficiently disambiguous as well as sufficiently rationally founded to be useful. The result is an overabundance of invalid static classification schemes that do not support the goals for which they are intended. In this work, we assessed the state of the field with regard to this issue,

characterized the properties that contribute to the construction of more valid static classification schemes, and identified two projects addressing the problem in productive ways. We further suggested other research avenues that have potential to make positive contributions and encourage new work in these areas.

Acknowledgements

Students in the graduate seminar on Forensic Software Engineering led by myself and John Knight at the University of Virginia in the Fall of 2002 provided a valuable environment for discussion of ideas that contributed to this paper. This work was partially funded by NASA grants NAG-1-2290 and NAG-02103, and NSF contract CCR-0205447.

References

1. Busse, D.K. *Cognitive Error Analysis*. Doctoral Dissertation, Department of Computing Science, University of Glasgow, 2002.
2. Greenwell, W.S., J.C. Knight, and E.A.Strunk. *Risk-Based Classification of Incidents.* 2003 Workshop on the Investigation and Reporting of Incidents and Accidents, 2003.
3. Hanks, K.S. and J.C. Knight. *In Search of Best Practices for the Use of Natural Language in the Development of High-Consequence Systems.* Supplement (FastAbstracts) to the Proceedings of the International Conference of Dependable Systems and Networks, 2002.
4. Hanks, K.S., J.C. Knight, and C.M. Holloway. *The Role of Natural Language in Accident Investigation and Reporting Guidelines.* Proceedings of the 2002 Workshop on the Investigation and Reporting of Incidents and Accidents, C. Johnson, ed., 2002.
5. Laboratory of Decision Making and Cognition, Columbia University. http://www.cpmc.columbia.edu/patel/ldmc/LDMCWebpage090802/FrontPage090402.htm, as on March 28, 2003.
6. National Aeronautics and Space Administration QS/Safety and Risk Management Division. *NASA Procedures and Guidelines for Mishap Reporting, Investigating, and Recordkeeping (NPG 8621.1),* 2000.
7. National Patient Safety Foundation. *End Stage Renal Disease Patient Safety Taxonomy.* http://www.esrdpatientsafety.com/taxonomy.html, as on March 28, 2003.
8. National Transportation Safety Board. *Accidents and Accident Rates by NTSB Classification, 1983 through 2002, for U.S. Air Carriers Operating Under 14 CFR 121.* http://www.ntsb.gov/aviation/Table2.htm, as on March 28, 2003.
9. Newell, A. and H. Simon. *Human Problem Solving*. Englewood Cliffs, NJ: Prentice Hall, 1972.
10. Perrow, C. *Normal Accidents: Living with High-Risk Technologies*. Princeton: Princeton University Press, 1999.
11. Petroski, H. *Design Paradigms: Case Histories of Error and Judgment in Engineering.* Cambridge: Cambridge University Press, 1994.
12. Rasmussen, J. *Information Processing and Human Machine Interaction*. Amsterdam: North Holland Press, 1986.
13. Reason, J. *Human Error*. Cambridge: Cambridge University Press, 1990.
14. Runciman, W.B., A. Sellen, R.K. Webb, J.A. Williamson, M. Currie, C. Morgan and W.J. Russell. *Errors, Incidents, and Accidents in Anaesthetic Practice.* Anaesthesia and Intensive Care 21(5): 506-519, 1993.
15. Transport NSW (New South Wales Department of Transport). *Signals Passed at Danger.* http://www.transport.nsw.gov.au/safety_reg/spad.html, as on March 28, 2003.

16. US Department of Transportation Federal Aviation Administration. *Aircraft Accident and Incident Notification, Investigation, and Reporting*. http://www2.faa.gov/avr/aai/TABL8020.htm, as on March 28, 2003.

Biography

Kimberly S. Wasson, University of Virginia; Charlottesville, Virginia, U.S.A.; telephone - +1.434.982.2225; fax - +1.434.982.2214; e-mail - ksh4q@cs.virginia.edu

Ms. Wasson is a doctoral candidate in Computer Science at the University of Virginia. Her primary interests include software requirements, software forensics, and linguistic and psychological issues affecting the quality of activities in both areas.

Newspaper and Online News Reporting of Major Accidents: Concorde AFR 4590 in The Times, The Sun and BBC Online

C.W. Johnson; Dept. of Computing Science, University of Glasgow, Glasgow, G12 9QQ.
Tel.: +44 141 330 6053, Fax: +44 141 330 4913; http://www.dcs.gla.ac.u/~johnson,
johnson@dcs.gla.ac.uk

Abstract

Many complaints have recently been made against the media reporting of major accidents (Johnson, 2003). It has been argued that undue emphasis is placed on identifying the immediate causes of any failure, including human error or technical failure, in the hours following an adverse event. In consequence, the public can be misinformed about the complex nature of many technological failures. The following pages present what is arguably the first detailed review of media coverage of a major accident. In particular, we consider the way in which a tabloid newspaper, a broadsheet and an Internet news service covered the loss of Concorde flight AFR4590 in July 2000. Our analysis yields some surprising results. The broadsheet speculates most about the causes of the incident, the tabloid publishes the least. The journalists and editorial staff on these new sources present very few direct hypotheses about the potential causes of this accident. In contrast, the majority of the speculation in the media is presented in the form of direct quotations from experts many of whom criticise undue speculation in the aftermath of such adverse events. This provides at least a partial explanation for the relative amount of speculative material in each of the publications that were studied. Experts may have been more inclined to speculate for the more prestigious broadsheet than they were for the mass-market tabloid publication. Alternatively, it can be argued that the editorial staff on the tabloid focussed their analysis more directly on the facts that were available in the aftermath of this accident.

Introduction

Investigatory authorities often have an ambivalent attitude towards the role of the media in the reporting of major accidents. Intense public interest in the course of any investigation must be balanced against the need to prevent undue or premature disclosure. This ambivalence is illustrated by the impact on the US National Transportation Safety Board of Section 515 of the Treasury and General Government Appropriations Act for 2001 (Public Law 106-554; H.R. 5658). This issued government-wide guidelines for maximizing the quality, objectivity, utility and integrity of information disseminated by Federal agencies. The NTSB (2002) acknowledged that; “The primary purpose of the NTSB is to promote safety improvements in the operations or oversight of public and private organizations, resulting in a safer transportation system in the United States. The primary audience of Safety Board products is persons, groups, or organizations that can bring about changes in transportation safety through action on the Board's safety recommendations. The Congress, industry, media, and public, who can influence the actions of the recommendation recipients, are also important audiences. The type of audience and the technical knowledge of the audience vary greatly, depending on the document's subject and the safety issues presented. The Safety Board does not intend its reports and recommendations to be read only by technicians and specialists in the transportation industry”. This wider role of the media in improving public safety forms a strong contrast with guidelines that govern the disclosure of information to the media in the immediate aftermath of an accident. The standard instructions from the senior investigator are that “The Safety Board will disseminate to the public all information regarding the accident [investigation], either through our Board Member, public affairs officer or me. We will hold regular briefings to the press. Please refrain from discussing the accident [investigation] in public, or giving information about it to the press. Any violation of

this request will be considered a serious infraction of Board rules". The NTSB (2000) argue that "This rule protects everyone. Typically, the NTSB conducts press briefings during the day and at night following the progress meeting. Only factual information -- that all the parties have heard -- is released. The NTSB does not speculate or give out unverified information. With all parties deferring to the Board to release information on the investigation, the team speaks in a coordinated, consistent and orderly manner. Through this procedure, competition for "spin" is thus minimized, and the maximum opportunity for coordination and cooperation among the parties is maintained".

Journalists often express a duty to inform the public about the causes of major accidents. This is eloquently expressed in the opening chapters of Downie and Kaiser's (2002) recent survey of 'American Journalism in Peril'. They argue that 'Communities are improved by aggressive, thorough coverage of important, if everyday, subjects like education, transportation, housing, work and recreation, government services and public safety'. For example, KHOU a local Houston television station played an important role in publicising a number of accidents involving Ford Explorers equipped with certain kinds of Firestone tires. The news coverage and federal investigations in 2000 led to the recall of millions of tires, "undoubtedly saving many lives". . The investigative role of the media is not restricted to KHOU. For example, both Le Parisien and the Times of London carried articles criticising the composition of the French Transport Ministry's investigation team into the loss of Concorde Flight AFR4590. Key individuals had investigated the crash of a French Air Inter Airbus in Alsace in 1992. Their report focused on the inexperience of the pilots, however, a subsequent court case identified the failure of cockpit instruments as a primary cause in this previous accident. Downie and Kaiser (2002) also point to the dangers of ill-informed coverage. They cite the example of journalists who were too eager to attribute the explosion of TWA Flight 800 to Islamic terrorists. They also argue that editorial policy can undermine good journalism; "If it bleeds, it leads is a self-mocking slogan among local television journalists, but also an accurate description of the reflex of television news directors…" Curtis' (1995) analysis of the New York Times' coverage of major airline accidents between 1978 and 1994 provides evidence to support this criticism of editorial policy. He used the Times' annual index of stories to argue that fatal events were also more likely to be reported as the number of fatalities increased. In particular, he argued that disproportionate coverage was devoted to 25 fatal airline events involving hijacks sabotage or military action. These events averaged 53 references each. The remaining 160 other fatal events averaged 7.2 references. The New York Times focused on events that occurred in the U.S. or that involved U.S. carriers.

<u>The Case Study: Concorde AFR 4590:</u> Curtis' review focussed on the coverage of many different incidents within a single newspaper. In contrast, the following pages focus on the reporting of a single incident. In particular, we focus on the articles that appeared in the aftermath of the Air France Concorde crash, flight AFR 4590. This decision is justified because the loss of AFR 4590 typifies the high-profile accidents that elicit considerable interest from the media. The official enquiry into this accident found that the front right tire of the left landing gear ran over a strip of metal shortly before rotation during takeoff from Charles de Gaulle Airport (BEA, 2002). The strip had fallen from another aircraft. Damage to the tire created debris that was thrown against the wing. The debris ruptured a fuel tank and a major fire broke out under the left wing. Problems appeared on engine 2 and for a brief period on engine 1 but the aircraft took off. The crew shut down engine 2, following an engine fire alarm. They noticed that the landing gear would not retract. The aircraft flew for around a minute but was unable to gain height or speed beyond 200 knots and 200 feet. Engine 1 lost thrust, the aircraft's angle of attack and bank increased sharply. The thrust on engines 3 and 4 fell suddenly and the aircraft crashed onto a hotel.

The Times, The Sun and BBC Online: The following pages analyse the coverage of the accident in two very different newspapers: The Times of London and The Sun. The Times is published in the large page area format associated with 'broadsheets'. It presents an authoritative, 'in-depth' analysis of news and current affairs and has a daily circulation of around 630,000 in August 2002. The Sun appears in the smaller 'tabloid' format. It presents news items but with a greater proportion of celebrity coverage and current affairs that The Times. The Sun enjoys daily sales of approximately 3,600,000. It is important to recognise, however, that newspapers are only one of several sources of news about incidents and accidents. In particular, there is a growing range of Internet based new services operated by organisations ranging from AOL-Time Warner, to the BBC and News International. At the time of the Concorde accident, most of these services were in their infancy. The BBC-online news service was in its second full year of operation. However, it was already the "most visited Internet content site in Europe" with the aim "to provide UK content in a market dominated by US material, and to act as a 'trusted guide'". The site aimed to cover more than 300 news items per day from around the globe. In the year before the Concorde accident BBC News Online attracted an average in excess of 3,000,000 hits per day, this resulted in an initial record of 120,600,000 million hits in March 2001. Although there are superficial similarities between newspapers, such as The Sun and The Times, and Internet news services, such as BBC Online, there are also numerous differences. For example, Internet services are not driven by publication and distribution deadlines. Stories can be edited on-line as more information becomes available 24-hours a day. Such differences complicate any comparative analysis between these news sources. For example, it is relatively easy to use newspapers to trace competing hypotheses about the causes of an accident by the careful reading of each successive edition. Things are less straightforward with Internet-based news services where any analysis must rely upon the timestamps associated with archives on particular servers. These times may only provide an indication of the last moment at which a story was edited and not the time when the document first appeared on a host website.

Quantitative Comparisons

This section presents a quantitative analysis of coverage about the crash of AFR 4590. It is quantitative in the sense that values are provided for the number of pages devoted to the subject in the days following the incident. Figures are also provided for the relative use of images, text and headlines in each of the three sources. The following sections provide a more subjective assessment of the different types of causal arguments that are used in the media as more evidence became available about the events leading to the accident.

Page Distributions for Coverage of the Accident: Figure 1 provides an overview of the coverage in The Sun, The Times and on BBC Online in the immediate aftermath of the loss of AFR 4590. It presents the total number of individual pages that contained references to the accident. This calculation is more complex than it might appear. As mentioned, previously, the analysis of the on-line resource depends upon access to an archive server. The total number of pages given in Figure 1 is the result of a query against the BBC archive using the term 'Concorde' restricted to the dates illustrated in the graph. The accuracy of the diagram, therefore, depends both on the precision and recall of the archive search engine. A second stage of analysis exhaustively analysed the returned documents to determine that only relevant articles were included. We did not, however, perform an exhaustive analysis of the several million pages that were excluded by the initial filtering process. Further complexity stems from the dynamic nature of on-line media. For example, several news items published on the 25th July were entitled 'Concorde Crashes in Paris'. The accident occurred shortly before 15:00GMT. The first of these pages was time stamped at 15:14 GMT and stated that "A Concorde jet flying to New York has crashed near

Paris Charles de Gaulle airport. The BBC correspondent in Paris say French TV is reporting that the aircraft crashed into a hotel shortly after take-off". A second page under the same title was time stamped at 15:42 and included an eye-witness account that the hotel was "totally in flames…I saw the Concorde go by with its left side engine on fire and crash a bit further away, about two minutes after taking off" (http://news.bbc.co.uk/1/hi/world/Europe/85093.stm). The initial story was revised five times over the day until the same headline was used on a more sustained piece that was finally published at 18:45. Figure 1 treats these pages as different news items even though it can be argued that one was a direct development of the other.

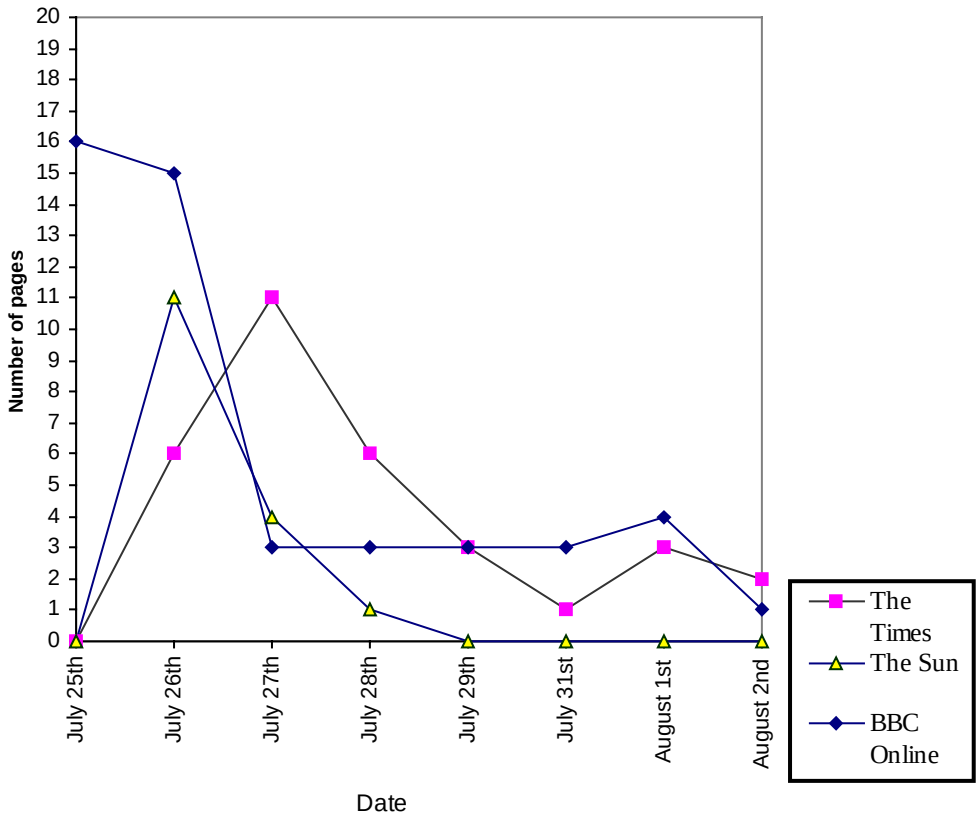


Figure 1 – Page Coverage of AFR 4590 by Date

A number of further issues complicate the development and interpretation of Figure 1. The Times and The Sun are both distributed across the UK. However, flexible production and distribution techniques were introduced across the newspaper industry during the 1980's and 1990's. One consequence of this has been that there are regional variations of national titles. These variations carry advertising and local news items that relate to the area in which the paper will be distributed. Figure 1 is based on a detailed analysis of the editions that were sold in Glasgow, Scotland in July and August 2000. The main news pages should be common across the distribution. It is possible, however, that some regional variations may affect our findings. This diagram excludes page totals from Sunday editions of The Times and The Sun. These papers are produced using different editorial teams, they have additional pages for more extended coverage and often repeat material that is published in the daily newspapers. The BBC Online

pages are collected from both the European and UK correspondents' contributions to the news service.

Some problems that frustrate the development of Figure 1 are common to both the newspapers and the web site. In particular, it can often be difficult to determine what exactly should be considered as 'news' and therefore be included within the page counts. BBC Online included several different categories of information. In particular, news coverage was distinguished from information about television programmes. By extending the scope of our search, it would be possible to increase the page count to include information about the BBC's wider broadcast coverage of this incident. This was not done and Figure 1 presents only the totals for pages that were produced by the BBC news staff. The Times includes a similar series of supplements, such as Times 2. In the aftermath of the Concorde accident these supplements included articles that considered media coverage by other European papers. Figure 1 includes these pages in the totals. This decision added 4 pages to The Times on July 27th, 2 pages on July 28th and a single page of coverage from a travel supplement on the 29th July.

Table 1 - BBC Online Coverage of AFR 4590 25th July 2000

Time Issued (GMT)	Title
01:18	The Cracks in Concorde
15:14	Concorde Crashes Near Paris (1)
15:42	Concorde Crashes Near Paris (2)
15:43	Q&A: Cracks in Concorde
15:50	Concorde Crashes Near Paris (3)
15:53	113 Killed in Concorde crash
15:55	Concorde facts and Figures
16:16	Concorde Crashes Near Paris (4)
16:25	Concorde Paris Crash Kills 113
16:33	Ageing Luxury Jet
17:02	Concorde: Loved by the Rich and Famous
17:15	Concorde 'Still the Safest'
17:56	Witnesses Describe Concorde 'Fireball'
18:45	Concorde Crashes Near Paris (5)
19:50	Concorde Kills 113 (2)
21:42	BA Suspends Concorde Flights
22:05	Germany Stunned by Concorde Crash

In spite of these caveats, a number of comments can be made about the media coverage based on Figure 1. An initial peak of interest can be observed in all three publications. This quickly declines over the following week. The way in which the coverage rises and then falls is different in each case. Both The Times and The Sun begin their coverage on the day after the accident. First reports were received on the afternoon of the 25th. The first national newspaper articles appeared on the morning of the 26th. The Sun devoted eleven pages of coverage on the 26th including many images from the scene of the crash and shortly before the accident occurred. The Times, in contrast, devoted most attention to the loss of AFR4590 on the 27th. It can be argued that this reflects an editorial policy of delaying publication until more facts are known in order to provide authoritative coverage. In contrast, BBC Online had the advantages of continuous publication over the Internet. As can be seen from Figure 1, most pages were devoted to this incident in the hours after the crash occurred. Table 1 provides further details of the headlines

that appeared for pages on the BBC Internet site in then hours after the crash. The steady accumulation of facts about the crash can be observed in these on-line archives in a manner that is not possible using daily newspaper publications where each edition summarises the information gleaned in the previous twenty-four hours. As mentioned, Figure 1 also illustrates the apparent decline in coverage across all three publications. This is most apparent in The Sun, which concentrated maximum coverage in the first edition after the crash. However, it is important to stress that much of the continuing coverage in The Times stemmed from readers' responses to previous articles rather than to stories produced by the papers' news staff. These letters account for a single page of coverage in The Times on August 1st and 2nd.

Figure 1 shows that BBC Online provided more sustained coverage than either newspaper. This is symptomatic of further differences between these forms of media. The Times' and The Sun's editors and journalists were faced with competing demands from other news stories for their finite column space. BBC Online did not face the same pressure of page limits as their more conventional counterparts. As a result, they continued to publish stories several weeks after the initial crash as, for example, Claude Gayssot the French Transport Minister coordinated the official response to the accident.

Relative Proportions of Text, Images and Headlines: Figure 1 arguably provides a false impression of the newspaper coverage in the aftermath of the Concorde accident. Although the BBC on-line pages were exclusively devoted to this topic, some of the newspaper pages contained very little information about the accident. As the week went on, full-page spreads were reduced to smaller articles. For example, page 13 was the only one to contain information about the accident in The Sun published on the 28th July. The total area of text devoted on that page was approximately 157 $cm^{2.}$ Figure 1 treats this in the same way as page 7 of The Times, which on the same day contained approximately 524 cm^2 of text at a smaller point size. Figure 2 presents a more detailed breakdown of media coverage following this accident. As mentioned before, BBC Online was able to publish its first articles within an hour of the crash. The newspaper response was delayed by publication schedules until the morning of the 26th. The additional detail in Figure 2 also illustrates important differences in the presentation of this incident. Both newspapers were able to use the delay before publication to acquire a large number of photographs taken during the last moments of the flight and in the subsequent operations to safeguard the crash site. The Sun's extensive use of these images, arguably, reflects the papers' format. However, it is important not to over simplify. Figure 2 also shows that The Times made extensive use of this photographic material. However, the proportion of images in The Times falls from 60% on the 26th to 45% on the 27th while the proportion of text devoted to the incident increases from 30% on the 26th. In contrast, BBC Online made less use of photographic images and correspondingly greater emphasis was placed on text-based reports. It could be argued that these photographs were not widely available at the time when BBC staff were beginning to assemble their first reports. However, the relatively high ratio of text to images is sustained into the 26th and beyond. This apparent difference between on-line and conventional press reporting can be explained by several important properties of the new Internet-based services. Firstly, many sites provide thumb-nail images that are embedded into the text of the new story. Readers can then choose to view higher-resolution images by selecting these thumb-nails. Hence, the ratio of text to images is, typically, quite different between screen space and the printed page. Secondly, there are well known differences in the readability of on-line versus printed text (Licorish). Most people will avoid reading long documents on CRT displays. Instead, they will either skim the prose, print it to read on paper or ignore it. In consequence, many of the on-line news provides impose guidelines on their journalists and editors so that few articles exceed 100-200 lines of prose. There is a conscious attempt to avoid unnecessary scrolling and reduce the demands imposed by on-line text.

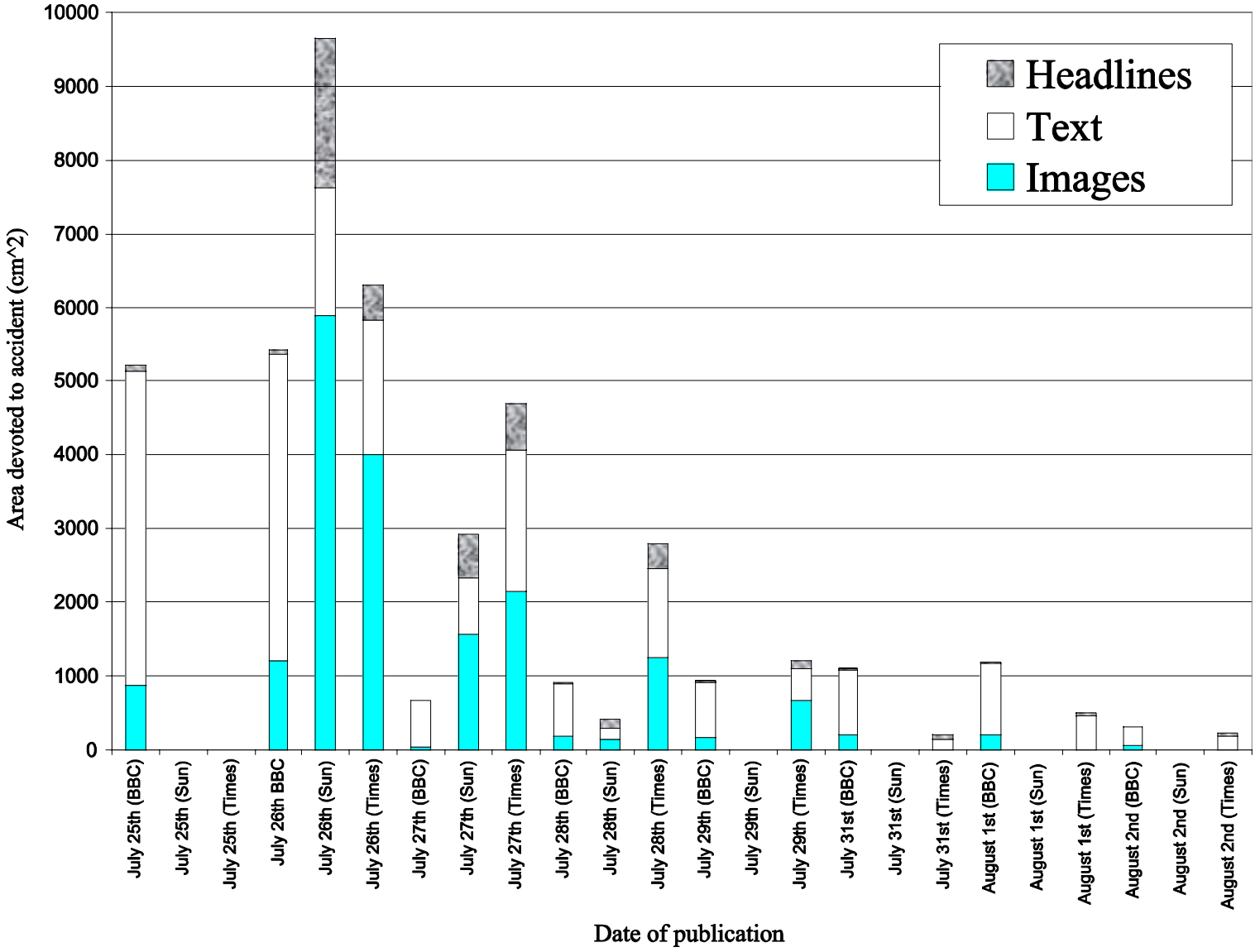


Figure 2 – Area of Text, Images and Headlines Devoted to AFR 4590 by Date

Table 2 - Area Devoted to AFR 4590 Excluding Supplements (cm^2)

	25th July			26th July			27th July			28th July		
	BBC	Sun	Times	BBC	Sun	Times	BBC	Sun	Times	BBC	Sun	Times
Text	4252	0	0	4158	1718	1829	622	760	1934	725	157	1196
Images	878	0	0	1208	5893	4000	40	1571	2146	180	144	1262
Headlines	85	0	0	66	2026	480	12	586	637	16.5	123	334

	29th July			31st July			1st August			2nd August		
	BBC	Sun	Times	BBC	Sun	Times	BBC	Sun	Times	BBC	Sun	Times
Text	768	0	444	884	0	154	962	0	452	250	0	194
Images	160	0	661	200	0	0	200	0	0	60	0	0
Headlines	15	0	108	19	0	48	25	0	59	5	0	34

Table 2 summarises the page areas devoted to the accident. It should be noted that the approximate total area in The Times' broadsheet format is 1,855 cm^2 and 945 cm^2 for The Sun's tabloid format. BBC Online provides a printable version of their articles with a total printable area of 416cm^2. These printed versions were used as a point of comparison between the on-line and newspaper sources. Further problems complicate any direct comparisons in terms of the total amount of text devoted by each source because The Times, The Sun and BBC Online use different point sizes and fonts. Taking the smallest point size used in each publication, a 40cm^2 area of text yields approximately 70 words in the printed version of BBC Online articles, 135 words in the 4cm column format of The Times and 170 words in the 5cm column format of The Sun. Matters are further complicated because different fonts and point sizes are used *within the same publication.* For example, The Sun uses 'strap lines' that lead the reader from the headline into the content of a story. These use a point size that is approximately midway between that of the headline and the main text. In Table 2, we have not accounted for the different word frequencies that are possible in the same area of prose at these different point sizes.

The problems that complicate the interpretation of Table 2 might be reduced if we could derive a word count for the Concorde articles using relatively simple computer-based tools. We could not, however, obtain complete electronic versions of the two newspapers that were being analysed. Even with access to the BBC Online documents it was difficult to derive accurate word counts. The task is complicated by the embedding of formatting commands, the use of style sheets and of inclusions from other pages of prose using frames. The only remaining solution is to perform a manual word count across the different media sources. The logistics of such an operation prevented us from exploiting this alternative. In contrast, the following pages look beyond the high-level statistics of this section. The intention is to focus more directly on the arguments that were presented in the media about the causes and the consequences of the Concorde accident. In particular, the intention is to identify the different hypotheses that were put forward about why the accident might have happened in the days following the loss of AFR 4590.

Qualitative Comparisons

The Bureau d'Enquêtes et d'Analyses pour la Sécurité de l'Aviation Civile enquiry into the accident argued that "front right tyre (tyre No 2) of the left landing gear ran over a strip of metal, which had fallen from another aircraft, and was damaged. Debris was thrown against the wing structure leading to a rupture of tank 5. A major fire, fuelled by the leak, broke out almost immediately under the left wing" (BEA, 2002). This information was not, however,

available to journalists and editors during the evening of the 25th and the morning of the 26th July. Instead, the immediate attention of all three sources focussed on previous reports about cracks having been found in the wings of the Concorde fleet. On Monday 24th July, British Airways staff had confirmed that hairline cracks had been discovered in the wing of all seven of its Concorde fleet. By coincidence, The Sun and BBC Online carried a series of articles on these 'problems' on the day of the crash. For example, the BBC reported, "one aircraft was grounded after a crack was found to have lengthened. BA was keen to stress the aircraft's exemplary safety record and the fact that Concorde clocks up a fraction of the flying hours amassed by sub-sonic planes" (BBC 848775.stm). This was published at 01:18 GMT on the 25th July. By 16:42 they were reporting, "The crash is the first of the supersonic jet built by Britain and France. It comes a day after British Airways confirmed that hairline cracks had been discovered in the wings of all seven of its Concorde fleet. The Concorde has been considered amongst the world's safest planes" (BBC 850903.stm). However, their account was also prescient in observing "its only scare came in 1979, when a bad landing blew out a plane's tyres. The incident led to a design modification".

Causal Hypotheses Changing Over Time: In the hours that followed the crash, the media revised their accounts. Experts argued that the cracks were unlikely to have played a significant role in the causes of the accident. By 17:15 on the 25th July, BBC Online were citing a former Concorde pilot who said that the cracks were "unlikely to have caused the French disaster" and by 19:50 "the Head of Air France said Tuesday's crash was linked to an engine problem and apparently had nothing to do with the cracks". The 21:42 update, however, quoted an aviation analyst as stating that "it is too early to speculate whether the plane has crashed because of this [the cracks]. The crash could have happened for a raft of reasons" (851057.stm). Over the following days, a number of diverse causal hypotheses were presented to the public. These ranged from age-related issues, including the possibility of metal fatigue, through to fan-blade separation within the engine or problems involving the maintenance of a thrust reverser immediately prior to take off. Table 3 provides an overview of how these different hypotheses appeared in the week following the accident.

Table 3 was obtained by an exhaustive reading of all of the material presented about Concorde in the three publications for the dates that are recorded in the top row of the diagram. A series of categories were devised from an initial read through and these are listed in the first column. The initial categories were then used to identify the causal hypotheses mentioned in each publication. However, this two stage classification process was not as straightforward as might be expected. In particular, several similar hypotheses were put forward with varying levels of detail. For instance, BBC Online on July 26th mentioned the possibility of a foreign object entering the intake of one of Concorde's engines. The Times on July 28th specifically mentions speculation about a bird strike contributing to the engine failure. The initial read through created the category of 'foreign object enters engine'. However, the more detailed hypothesis was retained in Table 3 from the second stage of the analysis to reflect the particular focus of The Times' article. Similarly, The Times contains speculation about the impact that staffing changes may have had on Concorde's maintenance before the crash while BBC Online stresses the relatively short time that was available to replace a thrust reverser that was found to be faulty immediately prior to take-off.

Table 3 - Potential Causal Factors by Date Discussed

	July 25th	**July 26th**	**July 27th**	**July 28th**	**July 29th**	**July 31st**	**Aug. 1st**	**Aug. 2nd**
The Sun								
Cracks in the wings	✓	✓	✓					
Age related issues (Including Metal fatigue)		✓						
Fan/turbine blade separation		✓		✓				
Uncontrolled release of fuel		✓		✓				
Thrust reverser			✓					
The Times								
Cracks in the wings	✓	✓						
Engine fire		✓						
Fan/turbine blade separation		✓	✓					
Failure in engine fire control system		✓						
Fractured fuel tank		✓	✓			✓	✓	✓
Hydraulic control failure		✓						
Terrorism		✓						
Human error		✓						
Tyre blow-out		✓		✓		✓	✓	✓
Age-related issues (Including Metal fatigue)		✓	✓	✓				
Thrust reverser			✓	✓				
Bird strike				✓				
Fuel line failure				✓		✓		
Maintenance staffing issues				✓				
Runway surveillance (foreign objects)						✓	✓	
After-burner ignition of fuel							✓	
BBC Online								
Cracks in the wings	✓	✓						
Engine fire	✓	✓		✓				
Other cause exacerbated by fuel load		✓						
Tire fragments damage engine				✓	✓	✓		
Tyre blow-out	✓	✓		✓	✓	✓	✓	
Thrust reverser		✓	✓		✓			
Foreign object enters engine		✓						
Fuel leak						✓	✓	
Lack of time for reverser maintenance			✓	✓				

Distribution of Different Forms of Causal Argument: The development of Table 3 was further complicated by the ambiguous manner in which causal hypotheses are often stated in the media. This was a particularly salient feature of the accounts of the Concorde crash. The Times, The Sun and BBC Online journalists rarely provided any direct speculation on the potential causes. When they did speculate, they were careful to stress the tentative nature of their suppositions. For example, The Times on the 26th July argued, "One possibility is that the fire control system in

the damaged engine failed to contain the problem, the fire damaged fuel lines, and power was lost in a second engine as the fire spread. A more remote possibility is that a fragment from a failed engine penetrated the aircraft's fuel tanks in the wing, causing a fire." (The Times, 26th July, p.5). Such direct speculation is, however, relatively rare. In contrast, the articles referred to previous problems, such as the cracks or tire bursts on landing, without making an explicit direct connection to the accident they were reporting. However, the reader is left to make an implicit connection between these previous incidents and potential causes of the loss of AFR4590. Similarly, potential causes are often raised and then immediately contradicted by other arguments. The Times on the 26th July also described how "the possibility of terrorism will be investigated, although Paris Charles de Gaulle has tightened up airport security in the last five years in the face of increased threats." (The Times, 26th July, p.5) and how "the most common single cause of major air accidents is human error, and the investigation teams will check on-board flight recorders and conversations between the pilot and air traffic controllers to find if there was any confusion in the last moments." (The Times, 26th July, p.5). Further rhetorical devices are used to avoid direct speculation. Arguably the most common is to rely upon experts to propose causal hypotheses. Again on the 26th, The Times describes how "Alan Smith, a former Concorde test pilot, said the most likely cause of the accident was a "catastrophic failure" of one of the plane's four engines. "It is possible that a turbine spun out from one engine and impacted upon the one next to it," he said.". (The Times, 26th July, p.1). There are further examples in the same edition, "John Guntripp, a former air crash investigator, said: "Even with two engines lost, the remaining two engines should have had more than sufficient power capable of taking the engine into a climb so what occurred was a very serious disruption of the aircraft's flying control. Conversations between the pilot and air traffic control will be recorded on one of the black boxes. The on-flight technical record will be checked to make sure that plane had been correctly serviced." (The Times, 26th July, p.3). Table 4 provides an overview of the distribution of these different forms of causal argument in The Times over the week following the accident. This was constructed by taking those sections of the articles that were identified as containing causal arguments in the first stage of developing Table 3. These paragraphs were then analysed to determine whether the causal argument was made 'directly' by the journalist as a claim about the loss of AFR4590. Each paragraph was also analysed to see whether it contradicted a possible causes, whether it contained direct expert testimony about a potential cause or whether it used indirect arguments about the causes of previous similar incidents. A single paragraph might be categorised under more than one of the rows in Table 4. For example, the following excerpt from The Sun would be classified as containing expert testimony contradicting a possible cause "BA's chief Concorde pilot, Mike Bannister said…"*These cracks, which the manufacturers have told us are non-safety related cause me no concern*. I have been aware of them for a little while and I have complete faith in BA's engineering and in the prudent steps they are taking to address a very small increase in the length of one of the cracks…" (26 July, p8). As can be seen in Table 4, quotations from experts provide most of the speculation about the causes of this accident. There is remarkably little direct speculation on the part of the journalists. It is important also to note the relatively large proportion of indirect arguments made in the hours following the crash by the Internet news service. This is unsurprising. Given the lack of any direct analysis, the journalists were forced to go back to report on the causes of previous incidents. The fourteen indirect causal factors mentioned on the 25th all related either to the microscopic wing cracks or to the tire burst on landing, mentioned above. The contradictions all relate to the wing cracks and none to the tire burst hypothesis. The direct causal hypotheses of the 28th and 29th July were substantially those confirmed in the BEA report, "The Concorde flight had been delayed for repairs to a thrust reverser, sparking early speculation that faulty work could have contributed to the disaster. But the investigators switched their focus to the burst tyre theory after shredded remains were found on the runway" (BBC 856606.stm).

Table 4 - Broad Overview of Causal Arguments in The Times

	July 25th	July 26th	July 27th	July 28th	July 29th	July 31st	Aug. 1st	Aug. 2nd
The Times								
Direct causal argument (X is a possible cause…)	0	2	5	1	0	1	0	0
Contradictions or caveats (X is unlikely as a cause…)	0	4	1	2	0	0	0	0
Indirect causal argument (X was a cause in the past…)	0	7	8	6	0	3	1	1
Expert quoted on cause (Y said X is possible cause…)	0	17	10	6	0	3	1	2
The Sun								
Direct causal argument (X is a possible cause…)	0	2	1	1	0	0	0	0
Contradictions or caveats (X is unlikely as a cause…)	0	2	0	1	0	0	0	0
Indirect causal argument (X was a cause in the past…)	0	6	4	0	0	0	0	0
Expert quoted on cause (Y said X is possible cause…)	0	10	7	3	0	0	0	0
BBC Online								
Direct causal argument (X is a possible cause…)	0	0	0	1	1	1	3	0
Contradictions or caveats (X is unlikely as a cause…)	4	5	0	1	0	1	2	0
Indirect causal argument (X was a cause in the past…)	14	7	2	3	5	4	3	0
Expert quoted on cause (Y said X is possible cause…)	13	10	0	5	9	2	4	0

The variety of causal arguments illustrated by Table 4 created particular problems in the construction of Table 3. It is often uncertain whether journalists and editors actually favour particular causal hypotheses when these different rhetorical devices are used. Any potential causes are usually introduced through expert quotations or are hedged by caveats and contradictory arguments. As a result, a tick in a cell of Table 3 denotes that a potential cause was mentioned in the pages of the associated publication on that date even if that cause may also have been questioned within the same article. This approach could be refined by introducing a system of ticks and crosses to indicate arguments for and against particular causal hypotheses. It can, however, be difficult to make definitive judgements about whether or not an argument supports or contradicts a potential cause. For example, The Sun on the 26th July quotes one expert as stating that "The stream of fire coming from the back of the plane is almost certainly burning fuel. Pilots who saw the burning plane said the flames spread to the second engine causing damage to that too. The explosion must have caused so much damage the fuel tanks cracked open and the flammable fuel spilled out…Concorde can fly with three engines no problem. But with just two there is real danger. At this point the plot must have lost control because the plane was moving too slowly to do anything. It is very likely the controls on the plane worked and the pilot was doing his best to avoid crashing into the hotel…" (26 July, p2). It is difficult to determine how many causal hypotheses are contained within such vernacular statements and whether one should

also assume that this account contains an implicit contradiction of previous hypotheses about the role of the cracks in the course of the accident.

Our use of the relatively simple 'ticks' in Table 3 is further justified by the need for independent validation of this subjective analysis. Another analyst should repeat the exercise and then some comparison should be made both between the causal categories and the identification of those categories in particular publications on a particular day. Unfortunately, it took 2-300 hours to complete the analysis that is summarised in Tables 3 and 4. This illustrates the need for greater research into the media reporting of major technological failures. In particular, we have previously described how software tools can in principle be used to automate much of this manual analysis using classification systems such as WordNet (Johnson, 2003). Having raised these caveats, it is possible to identify a number of tentative but potentially significant findings from this research.

A key finding from this research is that the tabloid Sun contains less speculation about the causes of the incident than the broadsheet Times. This is confirmed both in terms of the range of causal hypotheses that are considered, illustrated by Table 3, and by the number of paragraphs containing different forms of causal argument, illustrated by Table 4. A number of arguments can be put forward to explain this counter-intuitive observation. The official investigations provided little information in the immediate aftermath of the crash. The broadsheet was forced to speculate about alternate causes of the incident in order to sustain its analysis of the incident. It can also be argued that the higher profile and reputation of the broadsheet secured access to a larger range of experts who were more willing to be quoted in The Times than The Sun. It is difficult to find direct evidence to support this supposition. Table 4 does, however, illustrates that expert opinions form the major source of speculation for the broadsheet publication. The BBC Online site contains a wider range of causal hypotheses than The Sun but less than The Times. However, further analysis reveals that the Internet site devotes approximately 90 paragraphs to causal hypotheses while The Times provides just over 70. Hence BBC Online devoted greater space to a smaller range of causal arguments. This is not due to a greater level of detail in the Internet coverage. In contrast, it stems from the reiteration of the same hypotheses, as web pages are refined during a twenty-four hour period. For example, at 16:42 we find that "the crash is the first supersonic jet built by Britain and France. It comes a day after British Airways confirmed hairline cracks had been discovered in the wings of seven of the Concorde fleet." Exactly the same paragraph was included in the update issued at 16:53. The 17:16 page included the paragraph "A spokeswoman for Air France said all the passengers on board were Germans, on a special flight chartered by a German Tour operator. The crash comes a day after British Airways grounded one of its Concorde jets after small cracks were discovered in a number of the planes, although there's no suggestion the problem is linked to the crash".

Table 3 illustrates further differences between these media sources. The Sun focuses on fan-blade separation as a potential cause of the engine damage and fuel leak that led to the loss of AF 4590. In contrast, The Times and BBC Online consider a wider range of potential causes. However, both gradually converge on the possibility that a tire blowout may have fractured a fuel tank. This provides an important illustration of the way that information can be passed from the members of official investigations to the media. Even if this communication takes place through informal channels, it can effectively act to end the speculation that we see about alternate causes between the 26th and 29th July. It is also important to emphasise that individuals with appropriate skills and experience can also make prescient statements even if they are not part of an official investigation team. The Times identifies the afterburners as a potential ignition source in a letter from a fast-jet pilot in the RAF. Their comments pre-dated the BEA report that failed to

determine whether fuel ignition had occurred from a short-circuit in an electric harness close to the main landing gear or by fuel contact with hot sections of the engine reheating subsystem.

Mapping Causal Arguments Using Conclusion, Analysis, Evidence (CAE) Diagrams: Previous paragraphs have argued that there are important differences in the way that different section of the media handle the causal arguments that are made in the aftermath of major accidents. As we have seen, the broadsheet newspaper relies heavily on the use of expert opinion. The Online news service identifies fewer hypotheses but reiterates and refines them as stories are continually generated and updated. The tabloid has a more restricted palette of potential causes. Their coverage appears is not sustained in the same way that it is by the other news sources. One consequence of this is that most of their causal arguments rely on implicit references to the causes of previous incidents or existing safety concerns that may or may not have played a role in this particular incident. In all cases, there was remarkably little direct speculation about the events leading to the crash.

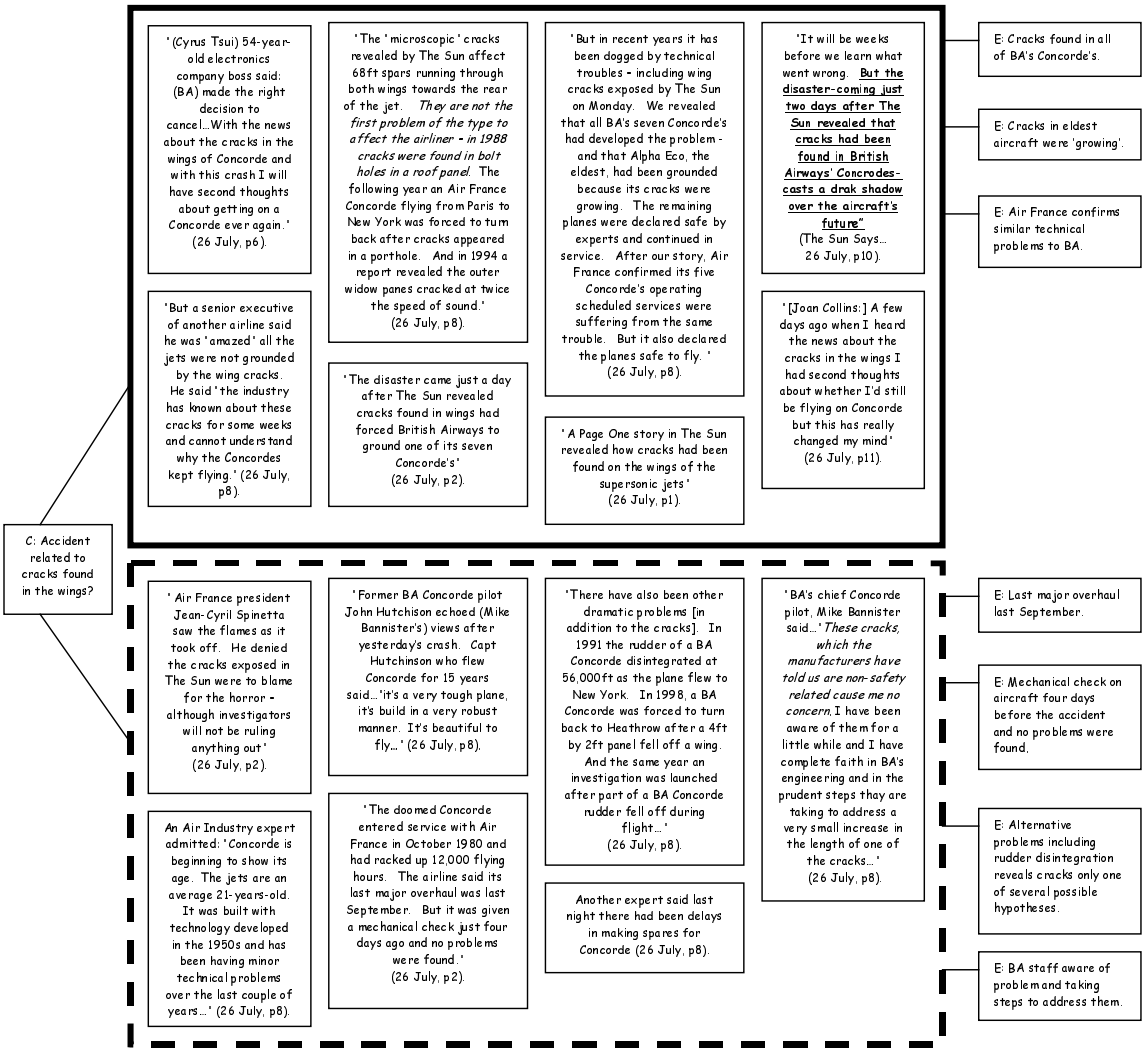


Figure 3 – Arguments Relating to the Presence of Cracks in the Wings, The Sun, July 26th

It is possible to probe beyond the high-level analyses presented in Tables 3 and 4. Conclusion, Analysis, Evidence (CAE) diagram provide means of mapping out the particular causal arguments that are presented about adverse events. Elsewhere we have used them to identify flaws in accident reports. For instance, inconsistencies can be identified where the same evidence is used both to support and weaken arguments about the cause of an accident. Similarly, an argument can be considered incomplete if it is not supported by links to the available evidence. Figure 3 provides an example of a CAE diagram applied to direct quotations from the Sun. In this case, it collates information about this potential cause that was presented on the day immediately following the accident. CAE diagrams provide a means of representing and reasoning about the arguments that are made in the aftermath of accidents and incidents (Johnson, 2003). The conclusion that cracks in the wing played a role in the accident is supported by a series of arguments that are represented by the large solid box on the top of Figure 3. For instance, the coverage on page 8 described how "'The microscopic' cracks revealed by The Sun affect 68ft spars running through both wings towards the rear of the jet. *They are not the first problem of the type to affect the airliner – in 1988 cracks were found in bolt holes in a roof panel.* The following year an Air France Concorde flying from Paris to New York was forced to turn back after cracks appeared in a porthole. And in 1994 a report revealed the outer widow panes cracked at twice the speed of sound." As can be seen, the arguments that support the involvement of the microscopic cracks are all indirect. The reader is left to infer that theis problem might have contributed to the loss of AFR4590 but this is not directly stated. In contrast, Figure 3 also illustrates arguments that weaken or contradict the involvement of these cracks. These arguments are represented in the dotted box in the lower part of the diagram. For example, page 8 describes how "There have also been other dramatic problems [in addition to the cracks]. In 1991 the rudder of a BA Concorde disintegrated at 56,000ft as the plane flew to New York. In 1998, a BA Concorde was forced to turn back to Heathrow after a 4ft by 2ft panel fell off a wing. And the same year an investigation was launched after part of a BA Concorde rudder fell off during flight…" The Sun also published more direct contradictions of this causal hypothesis, "Air France president Jean-Cyril Spinetta… denied the cracks exposed in The Sun were to blame for the horror – although investigators will not be ruling anything out". The evidence used in these different causal arguments is presented in the boxes on the far right of Figure 3.

Figure 4 extends the CAE analysis to illustrate the arguments that The Times made on the 26th July about the wing cracks. As can be seen, the CAE diagram immediately illustrates the more detailed analysis that is presented in the broadsheet. The same indirect forms of argument are used. For instance, on page five we read that "The investigation team...will be keen to know whether there is any connection between the crash and the recent discovery of small cracks in Concorde's wings. Both British Airways and Air France found the microscopic cracks within the last two months, but no aircraft was grounded until last week when the crack lengthened...both airlines insist that the cracks did not cause any safety fears." Figure 4 also illustrates the complexity of analysing the media coverage of causal arguments. The Times contains arguments that discount other causal hypotheses. For example, page 6 casts doubt on the potential terrorist threat to AFR 4590, "The possibility of terrorism will be investigated, although Paris Charles de Gaulle has tightened up airport security in the last five years in the face of increased threats." This argument has been included in the solid bounding box that supports the hypothesis about cracks in the wings. By attacking other causes, we can lend support to the remaining hypotheses.

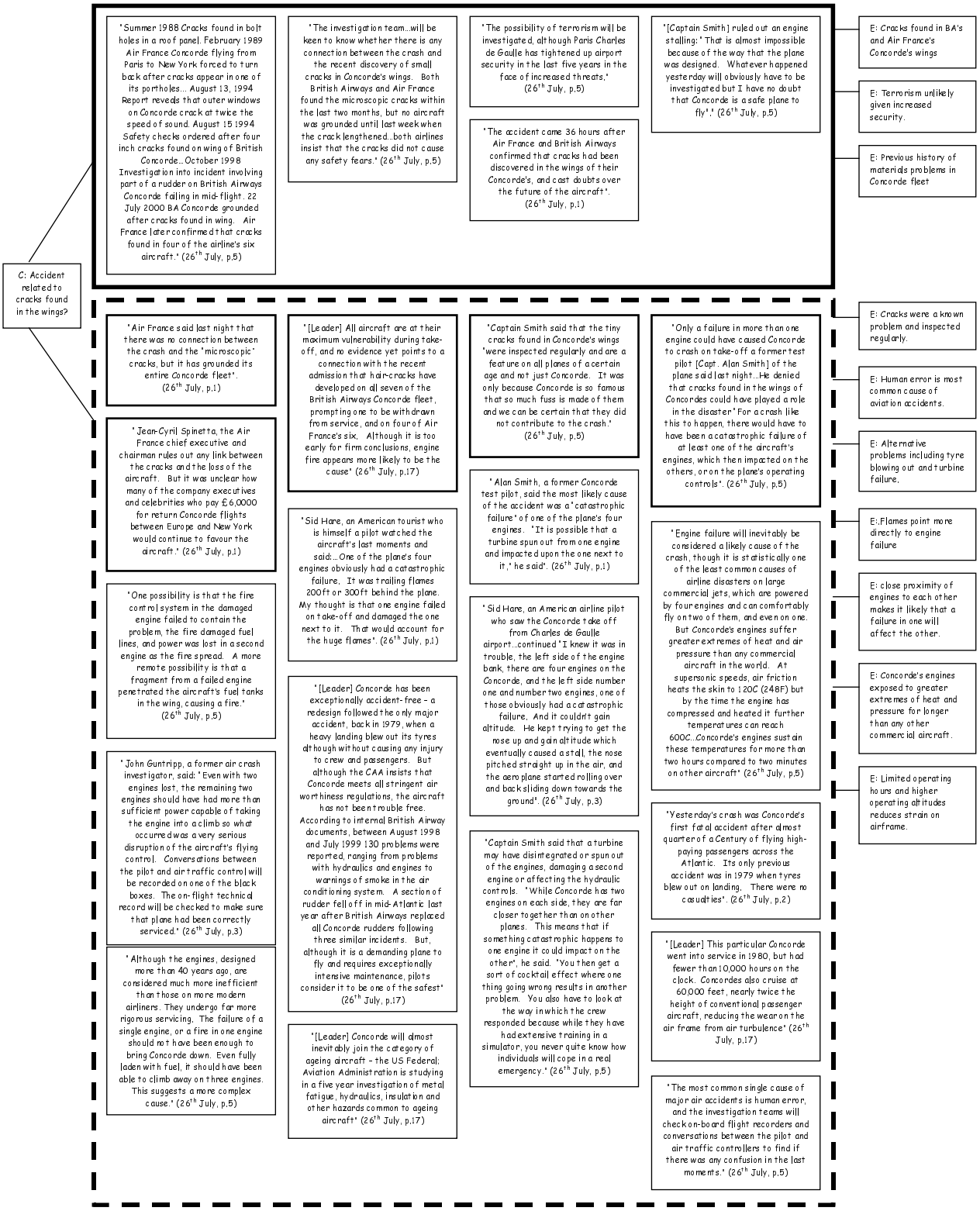


Figure 4 – Arguments Relating to the Presence of Cracks in the Wings, The Times, July 26th

Distinctions between indirect arguments suggesting that cracks might have been involved and arguments that cast doubt on other causes can be explicitly represented using more developed diagrammatical techniques. Figure 4 uses a heavier outline for arguments that directly contradict the role of the wing problems in the accident. Alternatively, the more complex argumentation diagrams developed by Toulmin (1999) might be used. In contrast, we retain the simpler CAE notation illustrated in Figures 3 and 4 partly because the solid and dotted bounding boxes provide an overview of the balance of arguments in the different publications. For instance, Figure 3 shows that the arguments in The Sun are almost equally divided for and against the role of the cracks in the accident. Figure 4 shows a greater degree of scepticism in The Times. This diagram also illustrates the prominent use of 'expert' opinion as a means of establishing causal hypotheses without journalists becoming drawn into more direct forms of speculation. For instance, on page 5 The Times cites the opinions of a former Concorde test pilot who "said that a turbine may have disintegrated or spun out of the engines, damaging a second engine or affecting the hydraulic controls. "While Concorde has two engines on each side, they are far closer together than on other planes. This means that if something catastrophic happens to one engine it could impact on the other", he said. "You then get a sort of cocktail effect where one thing going wrong results in another problem. You also have to look at the way in which the crew responded because while they have had extensive training in a simulator, you never quite know how individuals will cope in a real emergency." It is the opinions of these experts that cast the most doubt on the role of the cracks in the loss of AFR 4590. As mentioned previously, The Sun made less widespread use of such testimonies and this, in part, accounts for the greater emphasis that is placed on this causal hypothesis. The CAE diagrams in Figures 3 and 4 also illustrate further differences in the press coverage of this accident. The greater volume of prose and diversity of causal arguments in The Times do not rest on substantially more evidence than is presented in The Sun. This arguably underlines the dilemma facing broadsheet journalists. Their readers expect a more sustained analysis even though the staff must rely on information that is essentially similar to that available to their colleagues on mass-market titles.

Figure 5 shows how CAE diagrams can be extended to represent the competing causal hypotheses that emerged in the aftermath of the Concorde crash. In this case, the diagram represents arguments about the causes of the accident that appeared in articles on the BBC Online service between 25th and 29th July. This end date was chosen for convenience because it produced the largest CAE diagram that could be reproduced on a singe A4 page without paraphrasing the original arguments. As in previous diagrams, there is an element of subjectivity in the development of these figures. Only four causal hypotheses are mentioned. Other analysts might be able to identify other implicit arguments in the thousands of lines of prose that were published after this accident. The direct quotations in Figure 5 provide backing for the summary that is presented in Table 3. Initially attention focussed on the role played by the microscopic cracks in the wing. However, the BBC also referred to previous problems involving the tires on the day of the accident. The tire problems resurfaced some three days later when BEA investigators confirmed that debris had been found on the runway. The CAE diagram also illustrates the way in which some hypotheses were first raised and then dismissed. For instance, the repair to the thrust reversers was first mentioned on July 26th but was discredited by the 28th when 'investigators switched their focus to the burst tire theory".

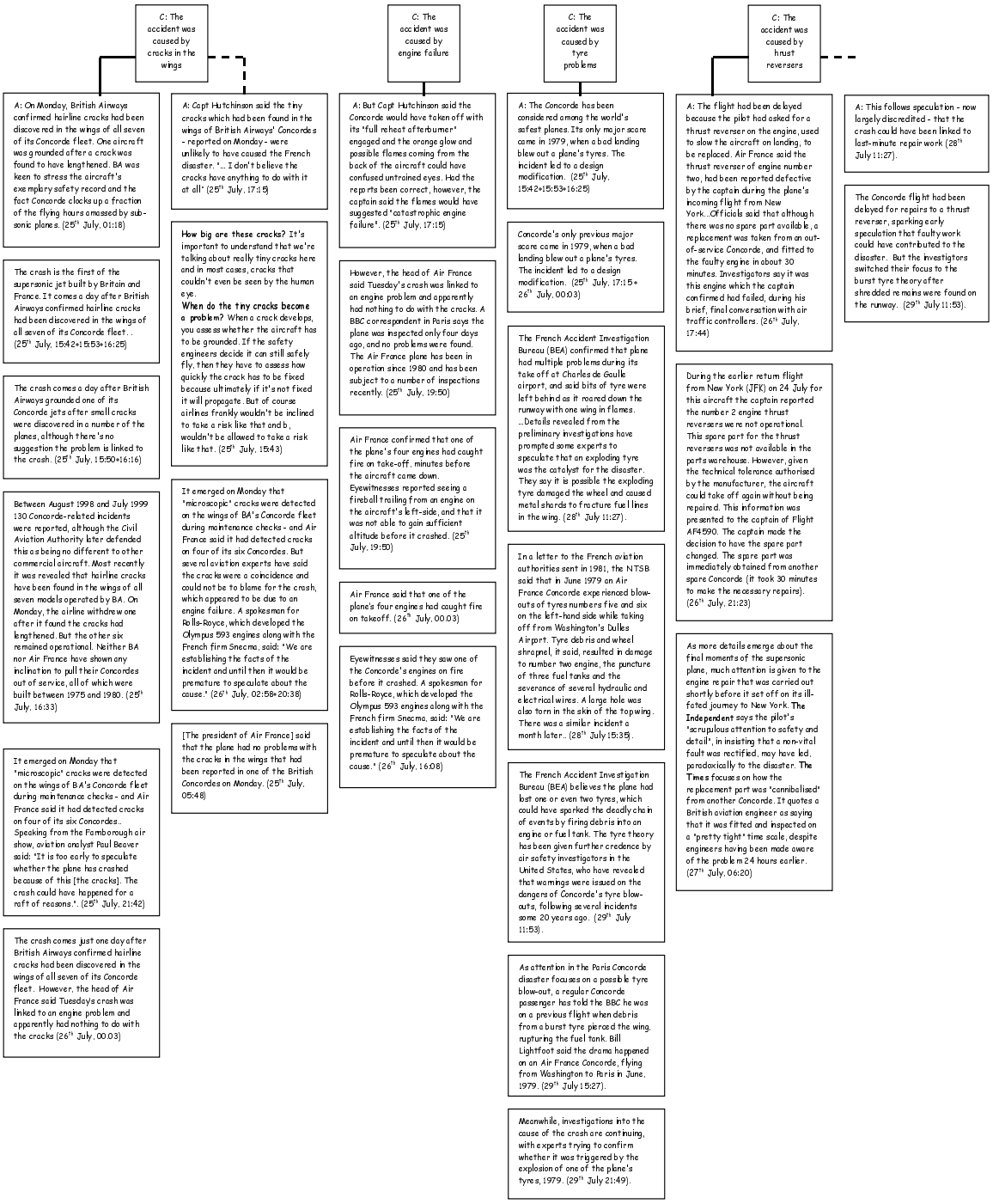


Figure 5 – Causal Hypotheses in the BBC Online Coverage of the Concorde Crash, 25th-29th July

Figure 5 illustrates indirect causal arguments of the form '(the accident) comes a day after British Airways confirmed hairline cracks had been discovered in the wings of all seven of its Concorde fleet". Contradictory arguments are illustrated by the observation that 'several aviation experts have said the cracks were a coincidence and could not be to blame for the crash, which appeared to be due to an engine failure.' As with the previous diagrams, however, there is little direct

speculation about the causes of the accident. Almost all of the hypotheses are put forward, or contradicted, by experts rather than be the journalists themselves. Although the media accounts speculate about the causes of the incident, they typically express the speculation in terms of direct quotes from safety professionals. They also offer alternative accounts that illustrate the uncertainty over these expert opinions. It is important to remember these insights when we condemn media speculation about the causes of accidents or incidents. They can best be thought of as a mirror that reflects the thoughts and opinions of the wider safety community.

Conclusions

This paper has analysed the reporting of the loss of Concorde AFR 4590 in three different news venues. We have compared articles published in a tabloid newspaper, The Sun, with a broadsheet, The Times of London and with an Internet based news service, BBC Online. Our study has focussed on coverage in the week following the accident. This decision was motivated by the sheer volume of material that was published in the aftermath of this adverse event. There have been very few previous studies of this type. Our results confirm some of the criticisms but challenge other assumptions that safety professionals have made about the media reporting of incidents and accidents (Johnson, 2003). In particular, we have noted the way in which an initial, high level of interest rapidly wanes as other new items compete for the finite column space of national newspapers. This effect is, however, less apparent in Internet news services that are free from some of the production and cost constraints that affect more traditional forms of publishing. There are other differences. Most notably, the Internet news service was able to start covering the accident almost within an hour of the crash occurring. The speed of response creates a dilemma for journalists who must provide copy about the adverse event at a time when little or nothing is known about what has taken place. We have also been able to identify important trends in the presentation of news coverage. Newspaper editors relied heavily on photographic images in their first editions following the accident. These images could provide an impression of what occurred without forcing journalists to provide detailed analysis of the potential causes. In the following days, readers were already familiar with these images and more information became available about the incident. In consequence, fewer images appeared and a greater proportion of the coverage was devoted to prose analysis of the potential causes.

The loss of AFR 4590 was deliberately chosen because it arguably represents the type of high-profile accident that would be most likely to encourage media speculation. This argument is strengthened by the way in which all of our news sources had covered the reported wing cracks on the morning of the 25th July. There is likely to have been an extremely strong temptation to directly link these warnings with the events that took place on the afternoon of the 25th. It is remarkable, therefore, that there was so little direct speculation in any of the sources that we examined. A further, paradoxical finding has been that the broadsheet account contains more speculation than the tabloid. We have argued that this is the result of a pressure to inform the readership about potential causes when little 'hard' information is available. Journalists seem to be aware of their dilemma and so speculation is, typically, presented in the form of direct quotes from experts and eyewitnesses.

A number of caveats must be made about this study. Firstly, we have only considered the media reaction in the week immediately following the accident. Further work is needed to analyse the subsequent reporting of the loss of AF 4590. Secondly, we focused on UK reaction. The nature of aviation accidents often creates media interest in several different countries. Most of the victims onboard AFR4590 were German. The aircraft was operated by a French company and crashed outside Paris. We are currently conducting a comparative study of the media reporting in these different countries. Thirdly, this paper has focused on two newspapers and an Internet

news service. More work is required to trace the causal analysis provided by broadcast services. Fortunately, the growth of publicly accessible digital archives has supported our work in this area. Having raised these caveats it is important to reiterate the central argument in this paper. Unless we understand the media reaction to major accidents then we will continue to repeat unjustified criticisms about their coverage of failures in safety-critical systems.

What does this study suggest for the regulatory and investigatory agencies that must address media concern in the aftermath of major accidents? This study has shown the importance of avoiding generalisation about the media's rush to speculate about the causes of an adverse event. The reliance on expert opinion suggests that greater attention might be paid to educating those safety professionals about the consequences of their speculation. The journalists already seem anxious to avoid direct speculation. Our study also revealed that speculation thrives in a vacuum. As soon as the BEA provided unofficial, indicative comments about the probable cause then all sources began to focus their coverage away from the more speculative comments. This is particularly apparent in the BBC Online coverage from the 29th July.

References

Bureau d'Enquêtes et d'Analyses pour la Sécurité de l'Aviation Civile (BEA), Accident on 25 July 2000 at La Patte d.Oie in Gonesse (95) to the Concorde registered F-BTSC operated by Air France, Report f-sc000725a. http://www.bea-fr.org/anglaise/actualite/concorde-en.htm, 2002.

T. Curtis, Airline Accidents and Media Bias: New York Times 1978-1994. Available from http://www.airsafe.com/nyt_bias.htm

L. Downie and R. Kaiser, The News About the News: American Journalism in Peril, Knopf, New York, 2002

C.W. Johnson, The Failure of Safety-Critical Systems: A Handbook of Accident and Incident Reporting, Springer Verlag, London, in press and to appear 2003.

NTSB Guidelines, Media Relations at Major Airlines and the National Transportation Safety Board, NTSB Office of Public Affairs,2000. Avaiable at http://www.wpntonline.com/tips/ntsb.html.

NTSB, Guidelines for Ensuring and Maximizing the Quality of Information Disseminated by the National Transportation Safety Board, 2002. Available on http://www.ntsb.gov/info/quality.htm.

E. Singer and P.M. Endreny, Reporting on Risk: How the Mass Media Portray Accidents, Diseases, Disasters, and Other Hazards, Russell Sage Foundation; May 1993.

S. Toulmin, The Uses of Argument, Cambridge University Press, Cambridge, 1999.

Automating Incident Analysis: A Challenge Paper

Fergus Toolan; Computer Science Dept., UCD; Dublin; Ireland
Joe Carthy; Computer Science Dept., UCD; Dublin; Ireland
Anne Drummond; Center for Safety and Health, UCD; Dublin; Ireland
John Dunnion; Computer Science Dept., UCD; Dublin; Ireland

Abstract

Government regulations have made the recording of incidents a legal requirement in many organisations. This data is kept for no purpose as most organisations never examine the information held in incident management systems. This paper examines possibilities for the automated analysis of these incident collections, hoping to learn from what has occurred previously to prevent future incidents and accidents.

Introduction

This paper discusses the area of automated incident analysis - the extraction of knowledge from an incident management system. We discuss some possible methods to use in this field and challenge the A.I. and I.R. communities to develop novel approaches to the automatic analysis of incidents.

This paper is structured as follows. We begin with an overview of Incident Management Systems and discuss some aspects of incident reporting and incident retrieval. Following this we introduce some possible methods to automate the incident analysis process. We then examine new approaches to evaluating the quality of the algorithms. It is hoped to distance ourselves from the requirement of having a domain expert present to decide the relevance of individual incidents. Next we discuss work relevant to this paper. Finally we summarise our contributions and mention some of the directions the authors are considering investigating.

Incident Reporting and Analysis

This section examines the two main components of Incident Management Systems: Incident Reporting and Incident Analysis. The following sections discuss these in detail and describe the current state of the art techniques in these areas.

Incident Reporting: Johnson [8] states that incident reporting schemes are increasingly being seen as a means of detecting and responding to failures before they develop into major accidents. He identifies seven benefits of using these schemes [7]. We focus on two of them:

- Incidents help to find out why accidents did not occur.
- The higher frequency of incidents permits quantitative analysis

These points suggest the need for automated incident analysis as according to the second point we have a large collection of incidents available to us. In most domains the incident collection is too large to examine manually and hence there is a need to automate the task.

The first point suggests the importance of automating the incident analysis task as it can prevent future accidents. These show the motivation for our work: the analysis of large collections of incidents to prevent future accidents.

However there are some problems with Incident Reporting Schemes that must be addressed. One of the big drawbacks in incident reporting is the feeling amongst staff that the system will be used to assign blame [13]. This in turn leads to people not submitting incident reports or submitting incorrect incident reports. This is a drawback to incident analysis be it manual or automated. If the reports are incorrect the analysis can never be accurate.

Root Causal Analysis: Root Causal Analysis is a means of identifying the causes of incidents / accidents using semi-rigorous methods to achieve this goal.

Safety through Organisational Learning (SOL) is an event analysis technique based on concepts of the socio-technical systems approach (STSA) and assumptions about accident causation [18] and Reason's Theory of Causation [14].

SOL uses a systemic view of safety, identifying five systems that must interact correctly to ensure safety of the system. The five are technology, individual, organisation, working group and organisational environment.

Why-Because Analysis: WBA [10] is one commonly used method of Root Causal Analysis. Indeed its use in industry is increasing with companies such as Siemens making use of the techniques in the analysis of accidents in the German rail industry [3]. WBA can be sub-divided into two important areas: 1) the creation of the WB Graph and 2) the verification of this graph.

The WB Graph allows us to identify the root causes of an incident / accident. It is in the area of WBA that Accident Investigators are most interested as it helps them to understand the causes of an accident. The verification step while much more complex gives a rigorous `proof' of the correctness of the WBG. It can be used to identify missing events / states in the graph. The ability to verify the graphs is the most striking difference between WBA and other Root Causal Analysis Techniques.

According to [9] the method used to generate the WBG is as follows

- **List:** List all events and states
- **Determine Causal Factors:** Use the Causal Factor test to determine the causal factors of an accident.

The second step in the technique is verification. This consists of a formal proof, the purpose of which is to ensure that

- the causal relations asserted by the WBG are correct, and
- that a sufficient amount of factors have been identified to provide enough support for the results of the WBG.

While the verification step is a difficult mathematical process, the power (and confidence) it gives are extremely important. The formal nature of the technique leads to the belief that it should be possible to automate this stage of the process.

Approaches to automated analysis

The major difficulties in automated incident analysis come from the standard representation of an incident that most Incident Management Systems use. These usually rely on small amounts of field-based information such as date, time, location...etc. and on large quantities of textual description in the form of witness reports, ATC reports...etc. The field-based information lends itself instantly to automated analysis however it is unlikely that using this information will result in any useful trends being discovered. The interesting information is often contained in the free text descriptions of the incidents.

To utilise this freetext information effectively for automated analysis we need to gain a better representation of this information. The challenge of the representation phase of automated analysis is to take the free text documents and extract the relevant information from them and store it in a format that can be used in pattern extraction algorithms.

This section is laid out as follows. Firstly we will look at the representation issues in incident reporting. This section also looks at how to translate from the human readable format that the report is in to a more machine-readable form. Following on from that we examine techniques for the automated analysis phase itself. Here we look at some established techniques such as Case Based Reasoning (CBR) and Data Mining. We also investigate new techniques in this area.

Representation: As already stated Incident Reporting schemes result in a combination of data types. We have numeric data such as the number of injuries or the number of people involved in an incident. We see temporal information in the form of dates, times and time intervals. Incident reporting schemes also result in textual information in two major formats: small text field information such as location and large textual narratives such as the description of the event. The heterogeneous nature of the information leads to difficulties in analysing the information. The techniques we discuss in the following sections are more suited to dealing with numeric fields and small textual fields. The major challenge lies in the fact that some of the most useful information is contained in the textual description fields of the report. This information is difficult to translate into a machine usable form.

Numerous techniques are possible to create a better representation of this information. We are mainly concentrating on WBA but any Root Causal Analysis technique may be used. We suggest that the WBG of the description may be the best representation to use. The graph representation is relatively straightforward to `mine' information from as will be seen later. However the transformation from textual to graphical representation may be quiet difficult to achieve.

Possibilities at this stage include a basic Information Retrieval approach using such techniques as stopword removal, stemming or n-gram extraction. However basic information retrieval loses some information vital to the incident analysis task - temporal information. We argue that it becomes impossible to accurately represent an incident without knowledge of the sequencing of the events. Take for example the process of starting a car. The driver ensures the car is in neutral, starts the engine, puts the car in first gear, releases the handbrake and drives away. A driver following these steps will successfully start the car. However, if instead the driver placed the car in gear, and tried to start the car it would stall. The driver in the second example is performing steps necessary to start a car but the sequencing means that they will never be able to do it.

This would suggest that Information Retrieval techniques need to be enhanced with something extra. Natural Language Processing provides some knowledge of sequencing of events. This

sequencing information is vital as with correct sequencing we narrow the possibilities for the root causal identification task. We know that a root cause of an event must occur before the event (or the state) it lead to and hence any event with no parent in the temporal hierarchy is a possible root cause.

To give a more formal representation of this we say if *A* and *B* are events in an incident description and if *A* occurs before *B* then *B* cannot be a root cause of the incident. This does not however, mean that *A* is a root cause only that *A* may be a root cause. When we have identified the WBG for the incident we are ready to combine this with the other fields in the incident management system relating to the incident in question and gain our final, complete description of the incident in a machine-readable form.

Automated Why-Because Analysis: The automation of the graph generation phase of WBA is one of the most complex tasks in automated incident analysis. The methods used to perform WBA do not lend themselves well to automatic execution by the computer. The first step, that of listing the events / states in the incident report is a very difficult natural language processing task. However with the nature of the aviation domain the task is semi-constrained making it much easier.

An example of this is the natural language problem of synonyms. Synonyms are two words that have the same meaning. An example of this in the aviation domain are the words **plane** and **aircraft**. However the technical nature of the domain has lead to each reporting scheme standardising this type of terminology. For instance the ASRS [19] dataset uses the term ACFT to mean airplane, plane or aircraft.

This constraint and others like it will hopefully allow more structured Information Retrieval techniques to be applied to the data rather than NLP techniques. Techniques such as stopword removal, stemming and n-gram extraction in conjunction with the most basic NLP techniques may be helpful in generating the events / states that occurred in the incident.

Case Based Reasoning: Case Based Reasoning (CBR) techniques can be used both to analyse incidents and to retrieve them from the incident management system. When used for retrieval they provide a `fuzzier' matching criterion than standard exact-match database queries [4]. In records as large as NASA's Aviation Safety Reporting System (ASRS) this is a vital facility as these records are too large to find exact matches. It can also be used for such things as judging the similarity between incident fields. For instance an incident involving a Boeing 737 would be more similar to one involving a Boeing 747 than a Cessna.

CBR techniques have been widely used to support a number of decision making tasks [7] such as faultfinding in the aviation domain. The decision necessary in incident analysis is the similarity of a new incident to others that occurred previously. These systems sometimes use a method known as Conversational CBR where the system has a set of questions it asks the user on encountering a new case. For instance in the technical support domain in the IT industry a question may be ``Does the monitor flicker?" Based on the answer the categorisation of the new problem gets one step closer to being correct.

NaCoDae (Navy Conversational Decision Aids Environment) uses Conversational CBR to discover incidents similar to a users query terms [1]. It uses a free text case representation which includes the appropriate solution to the problem if available. NaCoDae gradually refines the

users query through use of conversational techniques and hence overcomes inaccuracies in the users query.

Data Mining: Data Mining techniques are used to discover patterns in large Database collections [2]. In the domain of Incident Analysis they can be used to generate some basic patterns that are not necessarily obvious to the human user. In [16] we look at this area in more detail. This section gives an overview of Association Rule mining a very common technique in Data Mining. These focus on field-based information available in the dataset as oppossed to text-based information.

An Association Rule is a rule which implies certain association relationships amongst a set of objects in a database [2]. For instance, association rules could develop a set of symptoms associated with a disease or a set of items that commonly co-occur in a shopping basket.

Let L be a set of Literals (or items). An association rule is of the form $X \rightarrow Y$, where $X, Y \subseteq L$. The meaning of $X \rightarrow Y$ is that transactions that contain X tend to contain Y.

An association rule has two numeric terms associated with it namely its confidence and its support. An example rule is that "30% of transactions that contain beer also contain diapers: 2% of all transactions contain both of these items". We can define confidence and support in terms of this rule. The 30% value is the confidence. It is the number of transactions which contain X and also contain Y. The support value of 2% measures the percentage of occurrences of *both* X and Y in the set L. The problem is therefore to find all association rules that satisfy user-specified minimum support (S_{MIN}) and confidence (C_{MIN}).

Countless algorithms have been proposed for association rule mining. The best known is the *Apriori* algorithm which divides the problem into two separate parts

- Find combinations of items that have a transaction support above minimum support. These are frequent itemsets.
- Use frequent itemsets to generate the desired results. To do this assume $X \cup Y$ and $X \cup Y \cup Z$ are frequent itemsets then we can see if $(X \cup Y) \rightarrow Z$ holds by computing r, the ratio of $sup(X \cup Y \cup Z)$ to $sup(X \cup Y)$.
- If $r \geq C_{MIN}$ then $(X \cup Y) \rightarrow Z$ is a valid rule.

Many other algorithms exist for mining association rules. These include modifications to Breadth-First Search and Depth-First Search, and partition algorithms [5]. However, the more common solutions involve *Apriori* or new variations on the algorithm.

Classification Based Techniques: Classification based techniques are a standard Machine Learning technique that are used to decide how an item should be classified based on rules learned from a pre-classified set of items. Many forms of algorithms exist in this area. A recent application of basic classification techniques appeared in [6] using the horse racing domain.

Classification examines a set of data and generates a set of classification rules by which we can classify future data. This is very much in common with statistics and machine learning. In classification one develops a description or model for each class in a database based on the features present in a set of class-labeled training data [15] .

Various methods exist for mining classification rules [12]. The simplest forms are statistical algorithms such as linear models found in such packages as SAS or SPSS however, these don't scale very well. Another method is that of Neural Networks which try to copy the pattern matching ability of the human brain. Yet another commonly used technique is that of the nearest neighbour algorithm. This classifies each record in the dataset based on a combination of the classes of the *k* records most similar to it in a historical dataset. Another technique is rule induction which is the extraction of *if - then* rules from data based on statistical significance.

Graph Matching Algorithms: The techniques we have looked at so far for the analysis rely on the field-based information in the dataset and the new field-based representation of the narrative descriptions. However if root causal analysis is automated our final product contains the WBG (Why-Because Graph). If the automation of WBA is successful this will give us a new data format to try to analyse. The most likely candidates to be used in analysing this data is that of Graph Matching algorithms. Given that we are comparing two events Ev_1 and Ev_2 where $G_1 = (V_1, E_1)$ is the WBG for Ev_1 and $G_2 = (V_2, E_2)$ is the WBG for Ev_2. Let us define some notation for this: if G_1 is a subgraph of G_2 i.e. if $V_1 \subseteq V_2$ and $E_1 \subseteq E_2$ we write this as $G_1 \subset_G G_2$. If we wish to say that G_1 is a similar-subgraph of G_2 we write it as $G_1 \subset_{\overset{\approx}{G}} G_2$.

We are looking for a graph G_R where

$$G_R \subset_G G_1 \quad \text{and} \quad G_R \subset_G G_2$$

or

$$G_R \subset_{\overset{\approx}{G}} G_1 \quad \text{and} \quad G_R \subset_{\overset{\approx}{G}} G_2$$

The notion of a subgraph is a standard graph theoretic term. The above equations are saying that we are looking for a graph G_R which is a subgraph of both G_1 and G_2 or a similar-subgraph of G_1 and G_2. The notion of a similar-subgraph will be explained later.

Common Subgraphs Algorithm: The algorithm for extracting the common subgraphs of two graphs is shown in Figure 1. It begins by finding X_V and X_E the intersections of the Vertex set and the Edge set respectively. Then for every element of X_V it checks for members of X_E which are of the form (v_i, Z) where v_i is the current element of X_V and Z is any other element of X_V. When it finds these it adds v_i and Z to V_R - the vertex set of G_R - and adds the edge (v_i, Z) to the edge set of G_R.

This algorithm can extract all common subgraphs from two graphs with a complexity of $O(nm)$ where n is the cardinality of the intersection of the vertex sets, X_V, and m is the cardinality of the intersection of the edge sets, X_E.

Similar Subgraph's: In the domain of incident analysis the WBG is one of the best methods of visualising the sequence of events and the causal factors. When comparing WBG's we argue that the common subgraph is too restrictive a method to use. It is feasible to imagine a situation where an event, E_1 has a simple subgraph such as $A \rightarrow B \rightarrow C$ where $A \rightarrow B$ means A was a causal factor of B and event E_2 has as a subgraph $A \rightarrow C$. It is clear that these are not equivalent subgraphs however we argue that they are related.

Given 2 Graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$
Problem: Find G_R - the common subgraph
of $G_1 \wedge G_2$

Calculate $X_V : X_V = V_1 \cap V_2$
Calculate $X_E : X_E = E_1 \cap E_2$
$G_R := \emptyset$

$\forall\ v_i : v_i \in X_V$

$\forall\ e_j : e_j \in X_E \wedge e_j = (v_i, Z)$
AddVertex(G_R, v_i)
AddVertex(G_R, Z)
AddEdge(G_R, v_i, Z)
END
END

Figure 1 – Common Subgraph Algorithm

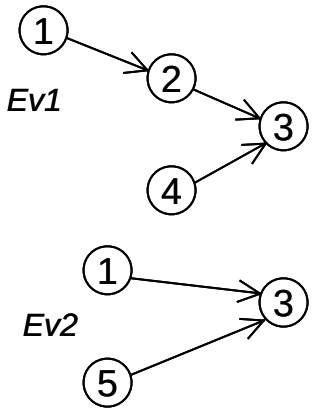


Figure 2 – WBG's for Ev_1 and Ev_2

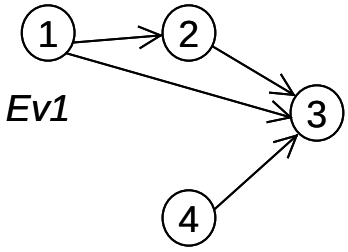


Figure 3 – Path Connected Ev_1

In both cases the event *A* lead either directly or indirectly to the event *C*. We argue there is a commonality here that the strict notion of common subgraphs would not allow us to exploit. We introduce the notion of similar subgraphs. Figure 2 shows two WBG's where states are represented by numerals.

Due to the nature of Causal graphs we can simplify the similar-subgraph detection problem by creating a ``path-connected" graph. The fact that WBG's are directed graphs allow us to connect the graphs along paths so every node in a maximal length path is connected in the direction of traversal. In Figure 2 Ev_2 has 2 maximal length paths of length 2, $1 \rightarrow 3$ and $5 \rightarrow 3$, while Ev_1 has 2 maximal length paths, one of length 2, $4 \rightarrow 3$, and one of length 3, $1 \rightarrow 2 \rightarrow 3$.

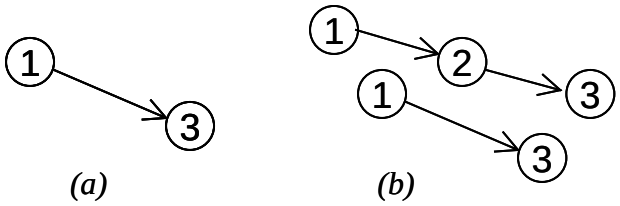


Figure 2 – (a) Common Subgraph (b) Similar Subgraphs

All maximal length paths of length 2 are by definition connected so in the above example the only unconnected maximal length path is $1 \rightarrow 2 \rightarrow 3$. Figure 3 shows the connected version of Ev_1. Creating the edge $1 \rightarrow 3$ is the only necessary step in creating the connected path. We can now use our common subgraph method to find common subgraphs and then editing them to get similar subgraphs. Figure 2(a) shows the common subgraph between Ev_2 and the connected Ev_1 while Figure 2(b) shows the similar subgraphs.

To generate the similar subgraphs from the common subgraph in Figure 2(a) we need to compare the common subgraph to the Ev_1 and Ev_2 graphs in Figure 2 and calculate any edges used in the common subgraph that were inserted at the connection phase. In this case we used $E(1,3)$ from Ev_1. We need to find the original path: this was $1 \rightarrow 2 \rightarrow 3$, and hence we have the simlar subgraphs shown in Figure 2(b).

Evaluation

Methods of evaluation can be quiet time consuming in these fields. A standard method of evaluation is to run the algorithms over a pre-classified set of examples and observe the level of accurracy it obtains. This is extremely difficult to achieve as it requires considerable person-hours from a domain expert to decide which incidents are related to each other.

However we have no valid method to avoid this method of evaluation. We can not evaluate the reliability of our algorithms without comparing them to a pre-classified set of incidents. This classification must be performed by a domain expert. It is possible to speed up the process by getting a domain expert to classify a small subset of the dataset and using a clustering algorithm to group other unclassified instances into these clusters.

Related Work

There is much work in the fields of Root Causal Analysis, CBR, Data Mining and Classification. However in terms of the application of CBR, Data Mining and Classification related to the Incident / Accident Analysis domains little has been done.

There is much work both in academia and industry on Root Causal Analysis. Numerous techniques are being used on a day-to-day basis. Such techniques include SOL [18], STAMP [11], and WBA [9] [10]. These techniques have been adopted (and altered) by industry and have been put to work in domains such as aviation [10] in the case of WBA, the nuclear power industry [18] in Germany uses SOL and the German Rail Industry where Siemens use a simplified version of WBA [3]. Work continues at pace both in refining the techniques used for Root Causal Analysis and the actual application of these techniques to discover the causes of accidents around the globe.

Data Mining has been applied to many domains from Horse Racing [6] to market basket data [2]. Both standard association rules [2] and more specialised techniques such as sequential pattern mining [15] are applicable in the incident analysis domain. Association rules are capable of generating commonly co-occurring items in a dataset while sequential patterns can predict common patterns of a sequential nature.

We are not altogether certain of the benefits of using association rule mining. Hobson-Shaw [6] found that in a dataset with many fields describing a single record that association rules were too general. For instance it would be feasible that the rule ``if the aircraft has an engine then it will crash" could be generated. Such a rule (while true based on the data examined i.e. every plane involved in an incident did have an engine) is too simplistic to be used in the real world. This leads us on to classification. Classification allowed [6] to discover more informative results.

Classification techniques such as Nearest Neighbour techniques and Decision Tree Learning Algorithms have been used regularly to classify items in a dataset into various categories. A new and interesting application of these techniques appeared recently in the form of classifications of winners from horse racing results in England [6].

As regards Case Based Reasoning, Cassidy et al [4] use CBR in a retrieval system for similar incidents in an incident management system. This can however, be viewed as a basic form of Incident Analysis in that the mere act of defining a similarity measure between two incidents is a method of analysing them. In this area the CBR methods outperformed the standard exact match methods of retrieval.

Conclusions

This paper has presented the challenge of automated analysis of incident report archives. Our work continues to focus on the area of automating Incident Analysis techniques such as WBA and the application of Data Mining techniques in the domain. In [16] we give some preliminary results from the application of Data Mining techniques to the domain. These by themselves are not extremely effective however we envisage a situation where these techniques used in conjunction with the methods of automating such techniques as WBA might prove to be extremely reliable.

References

[1] Aha, D. W., Breslow, L. A. & Maney, T. ``Conversational Case-Based Reasoning" in D.W. Aha & J.J. Daniels (Eds.) Case-Based Reasoning Integrations: Papers from the 1998 Workshop (Technical Report WS-98-15), 1998.

[2] Agrawal, R. & Srikant, R. ``Fast Algorithms for Mining Association Rules" in Proceedings of the 20th VLDB Conference, Santiago, Chile, 1994.

[3] Braband, J. & Brehmke, B. ``Application of Why-Because Graphs to Railway Near Misses" In Proceedings of Investigation and Reporting of Incidents and Accidents, Glasgow, Scotland, 2002.

[4] Cassidy, D., Carthy, J., Cullen, C., Hillary, P., McAdam, S., McGinty, L. & Sheppard, J. ``The Design and Implementation of an Incident Report Retrieval System" In Proceedings of

the 1st STarting AI Researchers Symposium, (STAIRS 2002) held in conjuction with ECAI 2002, Lyon, France, 2002.

[5] Hipp, J., Guntzer, U. & Nakhaeizadeh, G. ``Algorithms for Association Rule Mining - A General Survey and Comparison.'' SIGKDD Explorations, 2000.

[6] Hobson-Shaw, L. ``V. Wright - The Backer's Friend.'' Final Year Project Report submitted to the Computer Science Department, UCD, 2003.

[7] Johnson, C. ``Using Case-Based Reasoning to Support the Indexing and Retrieval of Incident Reports'' in Proceedings of European Safety and Reliability Conference (ESREL), Rotterdam, 2000.

[8] Johnson, C. ``Software Tools to Support Incident Reporting in Safety-Critical Systems'' in Journal of Safety Science, Elsevier, Vol 40, Number 9, 2002.

[9] Ladkin, P. B. ``A Quick Introduction Why-Because Analysis'' Available at http://www.rvs.uni-bielefeld.de/research/WBA/, 1999.

[10] Ladkin, P. B. ``Causal Analysis of Aircraft Accidents'' Invited Paper in Computer Safety, Reliability and Security, Proceedings of the 19th International Conference, SAFECOMP 2000, Lecture Notes in Computer Science No. 1943, Springer-Verlag, 2000.

[11] Levenson, N. ``A New Accident Model for Engineering Safer Systems'' in Proceedings of the MIT Engineering Systems Division Internal Symposium, May, 2002.

[12] Mitchell, T. M. ``Machine Learning'', McGraw-Hill, 1997

[13] O'Kane, M. ``Data Driven WebSite for Incident Management'' Final Year Project Report submitted to the Computer Science Department, UCD, 2003.

[14] Reason, J. ``Human Error'' Cambridge University Press, 1991.

[15] Toolan, F. & Kushmerick, N. ``Mining Web Logs for Personalized Site Maps'' in proceedings of the Workshop on Mining for Enhanced Web Search, International Conference on Web Information Systems Engineering (Singapore), 2002.

[16] Toolan, F. & Carthy, J. ``Data Mining Techniques applied to the Incident Analysis Domain'' IIRG Technical Report - TR-IIRG-03-01, 2003.

[17] Witten, I. H. & Eibe, F. ``Data Mining: Practical machine learning tools with Java implementations'' Morgan Kaufmann, San Francisco, 2000.

[18] ``Root Causal Analysis Literature Review'' Prepared by WS Atkins Consultants Limited for the Health and Safety Executive, 2001.

[19] Aviation Safety Reporting System (ASRS) http://asrs.arc.nasa.gov/

Biographies

Mr. Fergus Toolan, Intelligent Information Retrieval Group, Dept. of Computer Science, UCD, Dublin, Ireland.; telephone +353.1.716.2908; fax +353.1.269.7262; e-mail - fergus.toolan@ucd.ie.

Fergus Toolan is a PhD student in the Computer Science department in University College Dublin, Ireland. His research interests include Automated Incident / Accident Analysis, Data Mining and Information Retrieval.

Dr. Joe Carthy, Intelligent Information Retrieval Group, Dept. of Computer Science, UCD, Dublin, Ireland.; telephone +353.1.716.2481; fax +353.1.269.7262; e-mail - joe.carthy@ucd.ie.

Joe Carthy has been a lecturer in Computer Science at UCD since 1984. He has been actively involved in Information Retrieval research. His research interests include Topic Detection and Tracking, Incident Analysis and Retrieval and Lexical Chaining.

Ms. Anne Drummond, Centre for Safety and Health at Work, UCD, Dublin, Ireland.;
e-mail - anne.drummond@ucd.ie

Anne Drummond, MSc (Occupational Health), DipSHWW, RGN, RM, has worked in the UCD Centre for Safety and Health at Work since 1995 following a career in the healthcare sector. She is a College Lecturer, and Joint Academic Director of the Centre for Safety and Health at Work, which is an academic centre within the Faculty of Science.

Mr. John Dunnion, Intelligent Information Retrieval Group, Dept. of Computer Science, UCD, Dublin, Ireland.; telephone +353.1.716.2474; fax +353.1.269.7262; e-mail - john.dunnion@ucd.ie.

John Dunnion has been working in the Department of Computer Science in UCD since 1985 and has been on the academic staff since 1991. His principal research interests are in the area of intelligent text-based information retrieval.

Using Software Development Standards to Analyse Accidents Involving Electrical, Electronic or Programmable, Electronic Systems: The Blade Mill PLC Case Study

C.W. Johnson; Dept. of Computing Science, University of Glasgow, Glasgow, G12 9QQ, UK.
Tel.: +44 141 330 6053, johnson@dcs.gla.ac.uk

M. Bowell; Electrical & Control Systems Unit, Health and Safety Executive, Bootle, Merseyside, L20 3QZ, UK, Tel: +44 151 951 4064, Mark.Bowell@hse.gsi.gov.uk

Abstract

This paper presents the results of a project commissioned by the Electrical and Control Systems Unit of the UK Health and Safety Executive. The results of the project will be used to give guidance to operators and suppliers of electrical, electronic or programmable electronic systems (E/E/PES) in satisfying particular requirements of the Management of Health and Safety at Work Regulations 1999. The associated approved code of practice explains an obligation to '*adequately investigating the immediate and underlying causes of incidents and accidents to ensure that remedial action is taken, lessons are learnt and longer term objectives are introduced*'. There are relatively few techniques that might be used to investigate the underlying causes of E/E/PES related incidents. The following sections, therefore, introduce two techniques to support the investigation of this class of mishaps. One is based around flowcharts. These provide a series of questions to prompt investigators about the causal factors leading to an adverse event. Such a lightweight approach is appropriate for low consequence events. In contrast, the second technique involves additional documentation and analysis. It is, therefore, more appropriate for incidents that have greater potential consequences or a higher likelihood of recurrence. Events and Causal Factors (ECF) modeling is used together with a form of causal reasoning developed by the US Department of Energy (1992). The intention is that both the lightweight flowcharts and the more complex modeling techniques should help investigators to map causal factors back to the lifecycle phases and common requirements described in the IEC 61508 standard. This provides an important bridge from the products of mishap analysis to the design and operation of future systems. It is likely, however, that we will encounter incidents that cannot easily be attributed to lifecycle phases or common requirements in IEC 61508. Our work, therefore, offers important insights into the limitations of existing development standards. An implicit motivation in our work is to provide the feedback mechanisms that are necessary to improve the application of IEC 61508 and related standards such as DO-178B. A fatal injury in a gravel wash plant is used to illustrate this paper.

Introduction

The UK Health and Safety Executive's (HSE) mission is to ensure that risks to people's health and safety from work activities are properly controlled. An essential element of controlling risk is learning from past incidents and accidents – deciding the cause in each case and introducing new controls to reduce the risk of a repetition. To achieve its mission, HSE is supported by legal requirements, by approved codes of practice that interpret these requirements and by voluntary standards. The UK Health and Safety at Work Act 1974 places a legal duty on every company or organisation to reduce its risks "as far as is reasonably practicable". In other words, risks must be reduced until any further benefit is outweighed in gross disproportion by the effort required to obtain that benefit. In general, reasonably practicable measures are authoritatively defined in associated regulations and their approved codes of practice. They are also amplified through voluntary standards and guidance. The Management of Health and Safety at Work Regulations 1999 (HSE, 1999) require every employer to carry out a risk assessment, introduce the necessary

preventive and protective measures, and monitor these measures. The associated approved code of practice explains that monitoring includes:

1. *Adequately investigating the immediate and underlying causes of incidents and accidents to ensure that remedial action is taken, lessons are learnt and longer-term objectives are introduced.*

2. *It may be appropriate to record and analyse the results of monitoring activity, to identify any underlying themes or trends, which may not be apparent from looking at events in isolation.*

HSE is currently preparing general guidance material, possibly with supporting software tools, on how to investigate incidents and accidents. In parallel, HSE's Electrical and Control Systems Unit aims to produce cross-industry guidance on learning from incidents that specifically involve electrical and/or electronic and/or programmable electronic systems (E/E/PES). The terminology and conceptual framework for the E/E/PES technology specific work is taken from the international standard IEC 61508 (IEC 2000, 2003). This standard is applicable to all applications using this technology across all industry sectors, although the extent to which it applies will depend on other existing application and industry specific standards. IEC61508 includes requirements for developers and operators to learn from accidents and incidents (6.2.1-i of IEC 61508-1) and for suppliers to correct defects and report them to users (7.8.2.2 of IEC 61508-2). It does not give details on how to satisfy these requirements. In order to create some of the technical content necessary for HSE guidance, the Electrical and Control Systems Unit commissioned a multidisciplinary project on learning from incidents involving E/E/PE safety-related systems (HSE, 2003). The key stages of this project were to:

1. Evaluate existing schemes for analysing incidents, classifying data and generating lessons;
2. Consult users of existing schemes and potential users of HSE guidance;
3. Select and modify an existing scheme to integrate it with IEC 61508;
4. Test the new scheme using data from real incidents;
5. Present the scheme in the wider context of incident reporting, investigation and process improvement.

A companion paper describes the validation exercises in stages 4 and 5. This paper presents results from stages 2 and 3. The following section summarises the findings from our industry consultation into the reporting of E/E/PES related incidents. Subsequent sections introduce two new causal analysis techniques. A recent industrial accident described by the US Department of Labor's Mine Safety and Health Administration is used to illustrate the application of these techniques.

Industry Consultation: The development of our investigation techniques began with ten site visits to companies or organisations involved in the supply or operation of E/E/PES. Structured interviews were used to gather information about existing reporting procedures and mechanisms for disseminating any lessons learned from previous incidents. We were keen to identify perceived needs for incident reporting and investigation. The interviews were also intended to elicit any particular requirements for analysing E/E/PES related incidents. The industry sectors covered were pharmaceutical, nuclear, oil and gas, chemical process, marine, rail and machinery.

Roles included end users, designers, maintainers, procurers, assessors, system suppliers and component suppliers.

A number of key findings emerged from the consultation process. Comprehensive incident reporting and learning schemes that include the supply chain and information sharing are impeded by industry fragmentation. In particular, contracting out and the lack of continuity in the supply chain prevented any 'holistic' or 'systemic' approach. The user organisation's most significant technical influence over contractors is the standards used for project development. Many user organisations no longer have their own standards and instead reference international standards such as IEC 61508. Changes to these standards take many years. There are also competency and experience problems in most contract organisations. This applies both to operators and safety personnel. The majority of existing systems will not have been implemented using IEC 65108 as a design basis. There will be limited knowledge on the design history of such systems. Any guidance produced by HSE will need to be suitable for use with legacy systems. As might be expected, large end-user companies had the most sophisticated schemes especially where they are subject to the most regulation. End-user schemes were generic. In other words, they were not focused on E/E/PES. More than one company observed that the implementation of a more rigorous reporting scheme would increase the incident reporting rate, suggesting that there was previous under-reporting. However, they argued that if the scheme were successful then the increase in reporting rate might be offset by an anticipated reduction in the serious accident rate. Confidentiality could encourage reporting but most companies had non-confidential schemes. Management support and motivation is important for a successful scheme. This requires feedback to the reporters and investigators to show their activities are valued and acted upon.

Only a small fraction of reported incidents involved a special investigation of E/E/PES failure. For example, one company had 750 incidents per year, 6 were investigated in detail and only one involved this kind of special investigation. End user organisations often found it difficult to determine whether E/E/PES were implicated in an incident. Several causal analysis techniques were used. These included: timelines, event trees and checklists; a method similar to TRIPOD involving accident trees plus structured checklists (Johnson, 2003); event-based/event chain causal analysis (this company expressed dissatisfaction with their method, saying it did not get to the root causes very well); and ad-hoc approaches such as textual elaboration by designated experts. The E/E/PES suppliers did not use any specific method. In large companies we found up to four levels of internal incident enquiry depending on severity, e.g. trivial, local, formal investigation, formal enquiry, with different levels of investigation and different personnel at each level. Typically for large companies there were many thousands of trivial incidents per year but less than ten resulted in the most stringent type of enquiry. Some companies classified incidents according to type for subsequent monitoring and trend analysis. However, there was rarely any formal classification scheme of incident causes. The priority was to identify necessary changes in product, procedures or personnel competency. Recording of incidents, analyses and tracking of safety recommendations was quite sophisticated in some large companies and was implemented independently of other systems. However small companies tended to use existing QA systems for this purpose.

Some companies expressed concern about the costs of implementing any new scheme, for example in training and in writing new documentation and procedures. Also extensions to reporting might be a disincentive to both the reporters and the investigators if the process is too onerous. A new scheme should augment rather than replace existing systems, avoid technical language or jargon and communicate strengths and limitations clearly. Some companies had explicit mechanisms for reviewing and generalising incidents into recommendations. Experience was fed back into the design rules and business processes, and was often disseminated more

broadly to other sites, trade bodies and regulators. Tools such as databases, intranets, bulletin boards and e-mail aided dissemination. However this did not always succeed in changing company culture.

The Case Study Incident: This consultation process led to the development of two different analysis techniques. In order to illustrate the application of these tools, we introduce an incident that resulted in fatal injuries to a mechanic working in a gravel wash plant. This case study has been chosen because it is typical of the way in which incidents stem from the interaction between E/E/PES-related failures, hardware faults and management issues. The gravel wash plant cleaned and screened materials that were brought by truck from an off-site pit. The output from the operation was sold as part of a ready mix concrete business. The incident occurred inside a blade mill that was used to 'pre-condition' aggregates prior to wet screening. The mill consisted of two screws driven by two 40-horse power motors. The spiral grooves of each screw interlocked to help prepare the gravel. The motors were operated from a control center in a trailer about 30 meters from the mill. On the day of the incident, the mechanic and the wash plant foreman worked together to thaw frozen material inside the mill. They also intended to replace broken paddle tips and wearing shoes. The mechanic removed some sheets that had been placed on top of the mill to retain heat generated by a propane burner. This was being used to help thaw the frozen material. He then signalled to the foreman in the control center that he should start the mill motors in order to check that the blades were free. The motors started and so the foreman switched his attention to another task away from the mill. Before leaving, he switched the mill's start/stop buttons to the 'off' position. After completing his other task, the foreman returned to help carry out the necessary repairs on the mill paddles and shoes. However, the foreman was then called to assist an electrician who was working on a faulty circuit breaker. This had been tripping out after 10 to 15 minutes of operation. The electrician switched the breaker on and together with the foreman he watched it for several minutes without observing a trip. The electrician then turned it off and began to diagnose the problem. Meanwhile, the foreman returned to check on the mechanic. As he was leaving the control center, the foreman noticed that the two blade mill buttons were in the "run" position. He pushed them "off" and continued on to the mill where he found the mechanic entangled in the blades. Investigators determined that the mechanic had started the mill to clear some remaining frozen material after the foreman had left to work on his initial task away from the mill. The blades operated as the mechanic anticipated until the circuit breaker had tripped, before the electrician's inspection. For some reason, the mechanic then went back to work in the mill without shutting off any switches.

The faulty circuit breaker identified by the electrician controlled the power to several different systems including the control center lighting and the Programmable Logic Controller (PLC) that controlled the mill. A modification to the PLC approximately three months before the accident had resulted in power being unintentionally returned to components following a power failure, if their switches had been left in the "on" position. In consequence, the mill began operating when the breaker was reset during the troubleshooting by the foreman and the electrician. As can be seen, this incidents, stems from multiple causes. It was due to the failure to lock out the two-blade mill during the repairs. This stemmed from errors in the reprogramming of the PLC that allowed the automatic restart of equipment under control following a power trip. Further causes do not relate directly to the PLC. Power to the motor's circuit breakers was not locked out. No other measures were taken to prevent the equipment from becoming energized without the knowledge of the individuals working on it. In particular, the foreman was aware that the motor's circuits were not locked out while the electrician worked on the circuit breaker panel.

Root Causes of E/E/PES Related Incidents Under IEC61508

Several authors have argued that the root causes of complex, technological accidents often lie in decisions that were made months and years before the incident occurred (Leveson, 2002, Landkin & Loer, 1998). It is for this reason that our analytical techniques trace the causes of E/E/PES related accidents to problems in the development lifecycle. Latent causes can stem from the risk assessment process, during more detailed design, in implementation or in testing. Adverse events also often occur as a result of periodic maintenance, as was the case in the wash plant example. It is important also to recognise that other problems can affect several different stages of the lifecycle. For instance, poor documentation standards can carry problems forward from an initial risk analysis into implementation and beyond. Similarly, inadequate project management can undermine most development techniques. The causal analysis techniques presented in this paper, therefore, map the causes of E/E/PES related incidents to failures in the lifecycle stages and common process requirements in the IEC 61508 standard. This standard is one of several that could have been used (Johnson, 2003). The decision to adopt IEC 61508 is justified by its relatively widespread use in the process industries. HSE also recommended this general approach as the starting point for our work.

Table 1 provides a high-level classification of the potential problems that affect particular stages or are common to several different phases of the IEC 61508 lifecycle. The right column provides a reference to areas of the standard that provide additional detail about each requirement. The rows in this table will be used in the remainder of this report to provide a taxonomy or checklist of causal factors. As our analysis progresses we will identify which of these potential failures contributed to the particular causes of our case study. For example, an initial analysis of the wash plant example might argue that it stemmed from a modification failure. The verification and validation conducted after the reprogramming of the PLC failed to identify the particular failure mode that led to the incident. An important argument in this paper is that we must support investigators by providing tools that might help both to obtain and to justify such a causal analysis. The following pages, therefore, present two different techniques that can be used to map from accounts of an adverse event to the particular causes listed in Table 1.

Flow Charting Scheme: The flow-charting scheme provides a low cost technique for relatively low consequence incidents. Figures 1 and 2 present the current charts[1]. Analysis proceeds by asking a series of high-level questions about the nature of the E/E/PES-related incident. Investigators must determine whether or not the system correctly intervened to prevent a hazard, as might be the case in a near miss incident. If the answer is yes, then the investigator moves along the horizontal arrows. For instance, if the system intervened to address maintenance problems then they would follow the arrow in Figure 1 down to the associated table entry. By reading each cell in the column of the table indicated by the arrow, investigators can identify potential causes in the simplified stages of the IEC 61508 lifecycle. For instance, a maintenance failure might be due to problems in the risk assessment associated with the maintenance procedure or it might have been due to inadequate maintenance facilities and so on.

Table 1 - Taxonomy for Analyzing Computer Related Failures Under IEC 61508 (HSE, 2003).

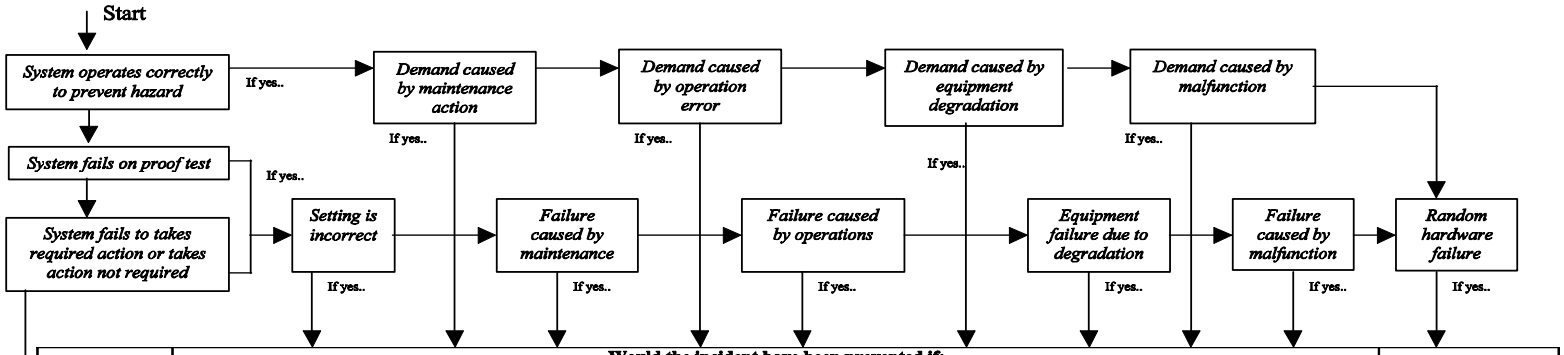
IEC 61508 Lifecycle phase	Detailed taxonomy	IEC 61508 ref
Concept Overall Scope Hazard & Risk Assessment	1. LTA Hazard identification 2. LTA Consequence and likelihood estimation	7.2,7.3,7.4
Overall Safety Requirements Allocation Planning of I & C, V, and O&M Realization	1. LTA specification 2. LTA selection of equipment 3. LTA design and development 4. LTA installation design 5. LTA maintenance facilities 6. LTA operations facilities	7.2 (2) 7.4.2.2 (2) 7.4 (2) 7.4.4/5 (2) 7.4.4.3(2), 7.4.5.2/3 (2) 7.4.5.1/3
Installation and commissioning	1. LTA installation 2. LTA commissioning	7.5 (2), 7.13.2.1/2, 7.13.2.3/4
Validation	1. LTA function testing 2. LTA discrepancies analysis 3. LTA validation techniques	7.7.2.1/2/3 (2) 7.7.2.5 (2) 7.7.2.7 (2)
Operation and maintenance	1. maintenance procedures not applied 2. maintenance procedures need improvement 3. operation procedures not applied 4. operations procedures need improvement 5. permit/hand over procedures 6. test interval not sufficient 7. maintenance procedures not impact assessed 8. operation procedures not assessed 9. LTA procedures to monitor system performance 10. LTA procedures applied to initiate modification in the event of systematic failures or vendor notification of faults 11. tools incorrectly selected or not applied correctly	7.7.2.1 7.6.2.2.1/2/3 (2) 7.6.2.1 7.6.2.2 7.6.2.1 7.6.2.1 7.6.2.4 (2) 7.6.2.4 (2) 7.6.2.1 (2) 7.8.2.2 (2), 7.16.2.2 7.6.2.1 (2)
Modification	1. impact analysis incorrect 2. LTA manufacturers information 3. full lifecycle not implemented 4. LTA verification and validation	7.8.2.1 (2) 7.8.2.2 (2) 7.8.2.3 (2) 7.8.2.4 (2)
IEC 61508 common requirements		
Competency	1. LTA operations competency 2. LTA maintenance competency 3. LTA modification competency	6.2.1 h 6.2.1 h 6.2.1 h
Lifecycle	1. LTA definition of operations accountabilities 2. LTA definition of maintenance accountabilities 3. LTA definition of modification accountabilities	7.1.4 7.1.4 7.1.4
Verification	1. LTA verification of operations 2. LTA verification of maintenance 3. LTA verification of modification	7.18.2, 7.9 (2) 7.18.2, 7.9 (2) 7.18.2, 7.9 (2)
Safety management	1. LTA safety culture 2. LTA safety audits 3. LTA management of suppliers	6.2.1 6.2.1 6.2.5
Documentation	1. documentation unclear or ambiguous 2. documentation incomplete 3. documentation not up to date	5.2.6 5.2.3 5.2.11
Functional safety assessment	1. LTA O & M assessment 2. modification assessment LTA 3. assessment incomplete 4. insufficient skills or independence in assessment team	8.2 8.2 8.2.3 8.2.11/12/13/14

Key: LTA is Less Than Adequate, IEC 61508 references are to Part 1 except as indicated by parentheses e.g. (2)

Investigators continue along the top horizontal line repeating the classification against the cells in the table in the same manner described for maintenance related incidents. In Figure 1, these address problems created by operator 'error', equipment damage and by equipment malfunctions. For some incidents, there will be failures identified by analyzing several of these different questions. A system may operate correctly to prevent a hazard although there may also be subsystem failures or operator interventions that initially fail to rectify the situation. In this case, analysts would focus on the top line in Figure 1 and the further line of analysis continued on Figure 2. The analysis might, therefore, help to identify many different causes on each pass through the flowchart. It is difficult to justify this exhaustive form of analysis for relatively minor incidents. In such cases, investigators may choose to stop once they have identified a potential cause from the flowcharts. Therefore, it is important that Safety Managers consider the order of questions in Figures 1 and 2. For instance, the current format asks whether maintenance issues potentially caused an incident before it elicits information about operator failures. This ordering can bias the analysis towards the causal factors that appear at the beginning of the flow chart. It is for this reason that we recommend a more sustained and exhaustive analysis so that investigators will consider the causes represented by subsequent entries. If this is not possible then safety managers should monitor the products of any causal analysis to identify the effects of ordering bias.

The flowcharts illustrated in Figures 1 and 2 have been validated against a series of case study incidents. These include the human factors related failure of a petrochemical production plant and a synchronisation incident in which redundant PLC pipelines shut down a floating production vessel (Johnson, 2003a). Each of the incidents that we have examined has helped to drive further refinements to the flowchart. We are currently conducting usability studies and validation exercises involving safety managers from across the process industries, including nuclear power generation and petrochemical production. These validation exercises also include participation from companies who supply and integrate E/E/PES applications. This is important because they are often called upon to identify the causes of mishaps that are reported by end-users. We also recognize that it may be necessary to tailor these flowcharts to the particular needs of an application domain. For instance, incidents involving E/E/PES in embedded systems are seldom caused by problems in the design and layout of graphical human computer interfaces. In contrast, user interface design has been at the heart of several recent incidents in petrochemical production (Johnson, 2003a). It should be stressed, however, that the flowcharts will become increasingly cumbersome as they are expanded to capture a growing range of potential causes. However, Figures 1 and 2 do illustrate our general approach to the analysis of less complex incidents and accidents.

It can be argued that such flowcharts constrain the identification of causal factors. They encourage very limited thinking about what contributes to adverse events. However, it is important to reiterate that we only advocate the use of this approach to support the initial analysis of low consequence, relatively simple mishaps. It is not intended to provide a panacea for the investigation of E/E/PES related incidents. It is, however, intended to provide a low cost approach that might replace the ad hoc techniques which are currently being used because many organisations lack the resources or motivation to use more complex approaches.

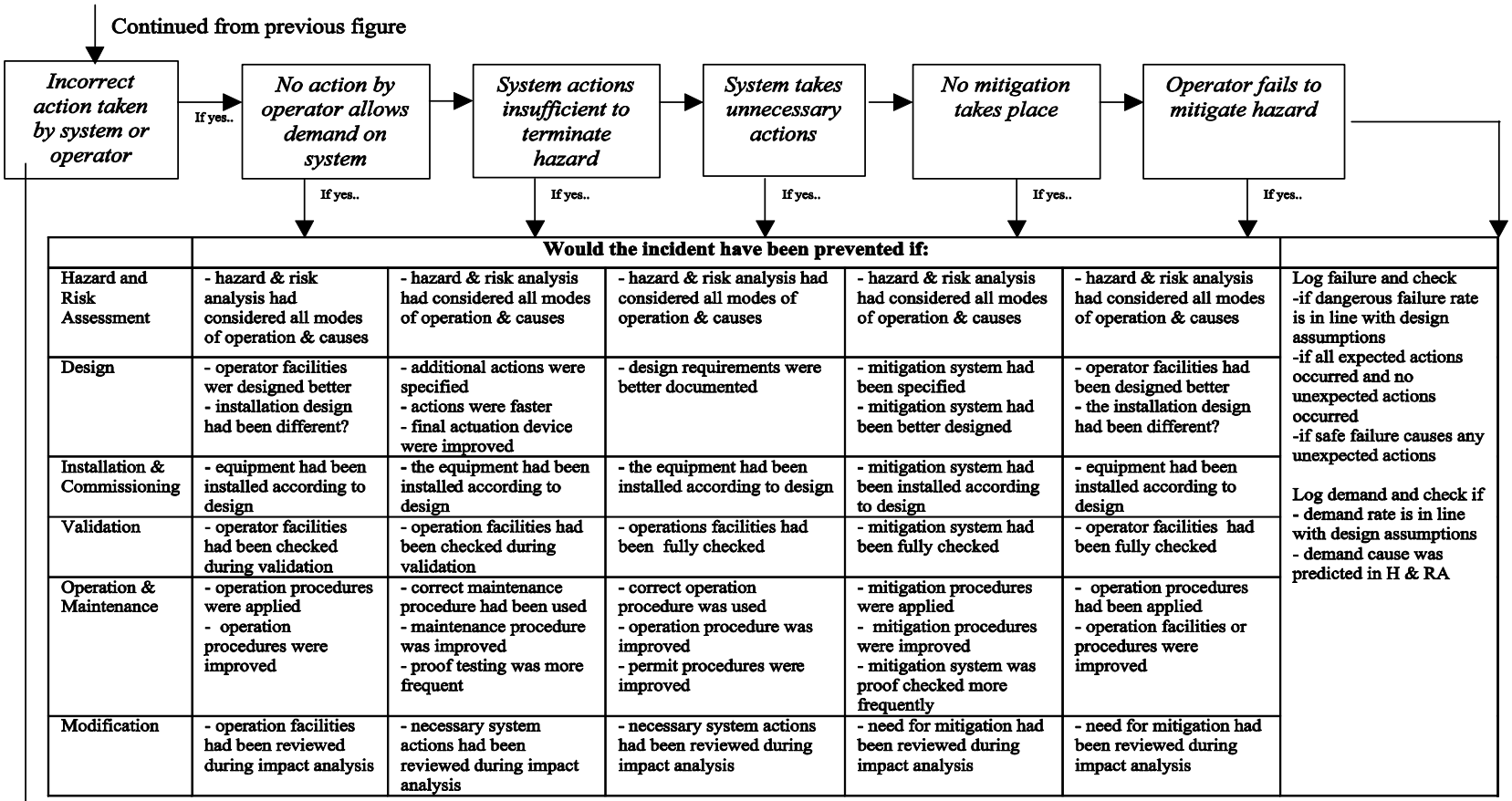


	Would the incident have been prevented if:					
Hazard and Risk Assessment	- hazard and risk analysis had considered all modes of operation and causes	- hazard and risk analysis had considered all modes of operation and causes	- hazard and risk analysis had considered all modes of operation and causes			Log failure and check -if dangerous failure rate is in line with design assumptions -if all expected actions occurred and no unexpected actions occurred -if safe failure causes any unexpected actions Log demand and check if -demand rate is in line with design assumptions -demand cause was predicted in H & RA
Design	- different equipment selected? - installation design different? - configuration was correct	- maintenance facilities had been designed adequately	- operations facilities had been designed correctly	- different equipment selected? - installation design had been different?	- different equipment selected? - the installation design had been different? -configuration was correct	
Installation & Commission	- the equipment had been installed according to design	- the maintenance facilities had been installed according to design	- the operations facilities had been installed according to design	- the equipment had been installed according to design	- the equipment had been installed according to design	
Validation	- the setting had been checked during validation	- maintenance facilities had been fully checked	- operations facilities had been fully checked	- equipment condition had been fully checked	- equipment condition had been fully checked	
Operation & Maintenance	- maintenance procedures were applied - maintenance procedures were improved - maintenance tools better - test interval was reduced	- correct maintenance procedure had been used - maintenance procedure was improved - permit procedures better	- correct operation procedure was used - operation procedure was improved - permit procedures improved	- maintenance procedures applied - maintenance proc. better - test interval was reduced - additional protection provided	- maintenance procedures were improved - maintenance tools improved - test interval was reduced - additional protection provided	
Modification	- setting had been reviewed during impact analysis	- maintenance facilities or procedures had been reviewed during impact analysis	- operation facilities or procedures had been reviewed during impact analysis	- equipment used or installation design has been reviewed during impact analysis	- equipment used or installation design has been reviewed during impact analysis	

	Would the incident have been prevented if:					
	Competency	Lifecycle	Verification	Safety management	Documentation	Safety assessment
Operation & Maintenance	- operation or maintenance staff were more competent	- responsibilities were defined better	- a better verification scheme had been in place	- safety culture was improved - audits were more frequent	- documentation was clear and sufficient	- assessment had been carried out on O&M phase
Modification	- modification carried out by more competent staff	- modification lifecycle was better defined	- a better verification scheme had been in place	- accountabilities better defined - suppliers not reviewed	- documentation updated	- assessment carried out on modification

Continued ...

Fig. 1 - High-Level Flow Chart to Support Causal Analysis of E/E/PES Related Incidents Using IEC 61508 Taxonomy [Continued in next figure] (HSE, 2003)



	Would the incident have been prevented if:					
Hazard and Risk Assessment	- hazard & risk analysis had considered all modes of operation & causes	- hazard & risk analysis had considered all modes of operation & causes	- hazard & risk analysis had considered all modes of operation & causes	- hazard & risk analysis had considered all modes of operation & causes	- hazard & risk analysis had considered all modes of operation & causes	Log failure and check -if dangerous failure rate is in line with design assumptions -if all expected actions occurred and no unexpected actions occurred -if safe failure causes any unexpected actions Log demand and check if - demand rate is in line with design assumptions - demand cause was predicted in H & RA
Design	- operator facilities wer designed better - installation design had been different?	- additional actions were specified - actions were faster - final actuation device were improved	- design requirements were better documented	- mitigation system had been specified - mitigation system had been better designed	- operator facilities had been designed better - the installation design had been different?	
Installation & Commissioning	- equipment had been installed according to design	- the equipment had been installed according to design	- the equipment had been installed according to design	- mitigation system had been installed according to design	- equipment had been installed according to design	
Validation	- operator facilities had been checked during validation	- operation facilities had been checked during validation	- operations facilities had been fully checked	- mitigation system had been fully checked	- operator facilities had been fully checked	
Operation & Maintenance	- operation procedures were applied - operation procedures were improved	- correct maintenance procedure had been used - maintenance procedure was improved - proof testing was more frequent	- correct operation procedure was used - operation procedure was improved - permit procedures were improved	- mitigation procedures were applied - mitigation procedures were improved - mitigation system was proof checked more frequently	- operation procedures had been applied - operation facilities or procedures were improved	
Modification	- operation facilities had been reviewed during impact analysis	- necessary system actions had been reviewed during impact analysis	- necessary system actions had been reviewed during impact analysis	- need for mitigation had been reviewed during impact analysis	- need for mitigation had been reviewed during impact analysis	

	Would the incident have been prevented if:					
	Competency	Lifecycle	Verification	Safety management	Documentation	Safety assessment
Operation & Maintenance	- operation or maintenance staff were more competent	- responsibilities were defined better	- a better verification scheme had been in place	- safety culture was improved - audits were more frequent	- documentation was clear and sufficient	- assessment had been carried out on O&M phase
Modification	- modification had been carried out by more competent staff	- modification lifecycle was better defined	- a better verification scheme had been in place	- accountabilities were better defined - suppliers not reviewed	- documentation had been updated	- assessment had been carried out on modification

Fig. 2 - High-Level Flow Chart to Support Causal Analysis of E/E/PES Related Incidents Using IEC 61508 Taxonomy (HSE, 2003).

Our case study, as with many E/E/PES related incidents, stems from multiple causes. It was due to the failure to lock out the two-blade mill during the repair operation. This, in turn, was due to errors in the reprogramming of the PLC. This allowed the automatic restart of equipment under control following a power trip. There are further causes that do not relate directly to the PLC. For example, the power to the motor's circuit breakers was not locked out. No other measures were taken to prevent the equipment from becoming energized without the knowledge of the individuals working on it. In particular, the foreman was aware that the motor's circuits were not locked out while the electrician worked on the circuit breaker panel. Several requirements or lifecycle activities might have prevented this incident from occurring in the manner described. Table 2 illustrates one means of documenting the products of any flowchart analysis. Immediate events that are identified in incident reporting forms are related back to failures in the lifecycle stages and common requirements of IEC 61508. This allocation process is guided by the questions in Figures 1 and 2. Errors in the reprogramming were due to an inadequate hazard analysis. This failed to identify the potential failure modes associated with allowing the automated restart of equipment under control following a power trip.

Table 2 - Abridged IEC 61508 Flowchart Causal Summary for the Case Study

Causal Event	**IEC 61508 Classification**	**Route through flow chart**	**Rationale**
PLC allows automatic restart of equipment following power trip	**Hazard and risk assessment**	System fails to take required action -> Failure caused by maintenance -> Hazard and risk analysis had not considered all modes of operation.	The reprogramming of the PLC allowed for a situation in which equipment was automatically restarted following a power trip. Reprogramming is likely to have prevented a restart without operator intervention had this potential hazard been recognised. (Note: if there were evidence that this hazard had been considered during the reprogramming then the causal analysis might have focussed more on validation to ensure that the PLC prevented the automated restart hazard.)
Failure to warn mechanic that power circuits not locked out during maintenance on circuit breaker.	**Operation and maintenance**	System fails to take required action -> Failure caused by maintenance -> Accident would have been avoided if maintenance procedure were improved.	On-site investigators argued that the foreman was aware of the relationship between the circuit breakers and the mill. The incident might have been avoided if they had followed a documented maintenance procedure or permission to work scheme that would have locked out all equipment affected by the maintenance on the circuit breakers.

Event & Causal Factor Analysis: Table 2 provides a relatively high-level form of causal analysis. Such techniques are appropriate for low consequence incidents. They might also be used during the initial stages of an investigation. It is unlikely that the flowcharts of Figures 1 and 2 will prove sufficient for more serious or complex incidents. The following section, therefore, presents a more sophisticated approach. It also enables investigators to map the causes of an adverse event to failures in the development lifecycle.

First Stage: Information Elicitation and ECF Modelling

The first stage in our more complex, causal analysis technique is to map out the events and conditions that led to the incident. Figure 3 uses a simplified form of the Events and Causal Factors (ECF) diagrams that were developed for the US Department of Energy (1992). Rectangles represent events. Ovals represent the conditions that make those events more likely.

The diamond shape represents the outcome of the E/E/PES related mishap. This technique was chosen after extensive discussions with individuals involved in the development and application of the IEC61508 standard and after consultation with HSE representatives. This does not, however, imply that ECF modeling is the only technique that we might have used. Leveson (2002) has recently challenged the utility of event based modeling techniques. She has argued that greater attention should be paid to the constraints that hold between system components. For example, by focusing on the actions of the foreman we might overlook the key requirement that blade motors are not automatically restarted on power-up. Leveson's alternative techniques do, however, rely upon an initial reconstruction. The subsequent stages of her STAMP method also have much in common with the approach in this paper. Rather than focusing on the violation of development lifecycle requirements, Leveson identifies more general failures to satisfy the constraints that should hold between system components. Hence our approach focuses more on problems in the development process rather than deficiencies in the final system. This is justified because the same development processes may have been used well beyond the boundary of the particular system involved in a particular incident. A further difference stems from our insistence that the investigation technique should inform the subsequent refinement of safety-critical development standards, such as IEC 61508.

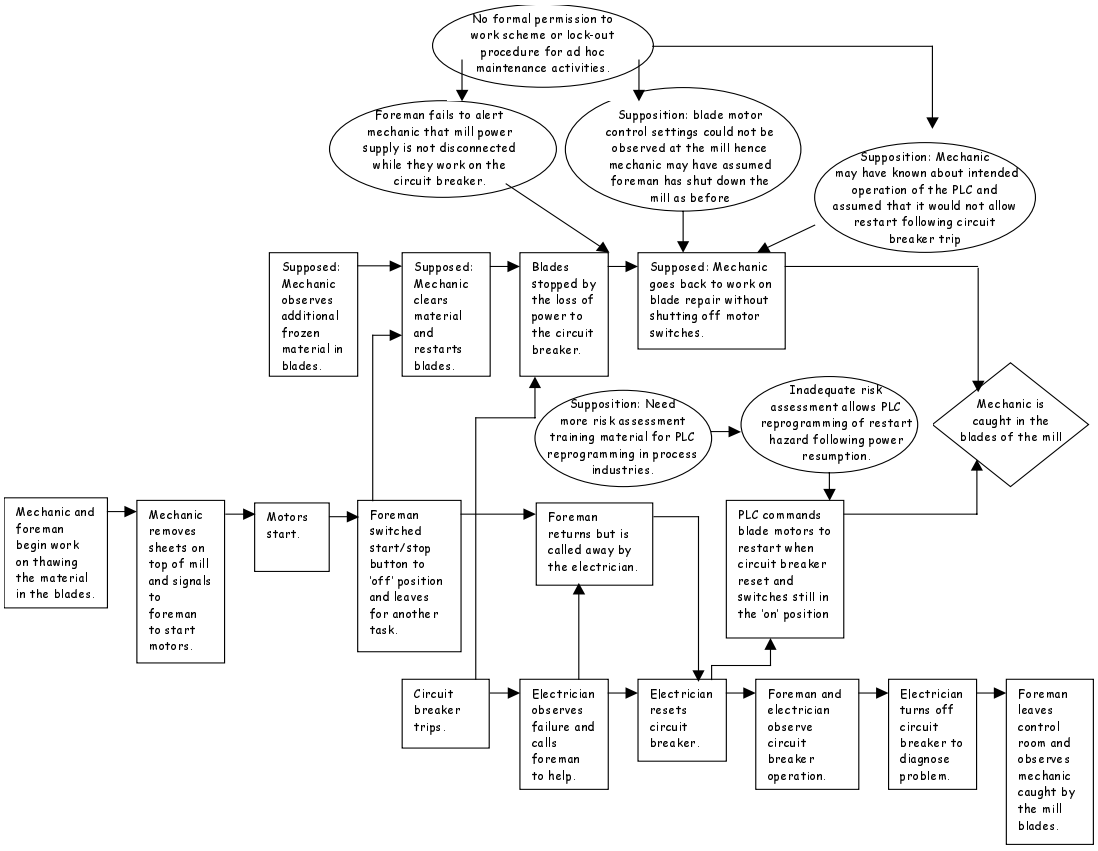


Fig. 3 - ECF Diagrams Including Developer/System Integrator Information

Figure 3 uses the ECF notation to represent the events and conditions that ultimately lead to the operation of the mill blades while the mechanic was repairing the mill. As can be seen, key events include the mechanic return to work on the blade repair without shutting off the motor

switches and the electrician's decision to reset the circuit breakers. Conditions include an 'inadequate risk assessment for maintenance procedures on the PLC update, allows restart hazard following power resumption'. The ECF chart provides a common focus for multi-party investigations. The development of this diagram continues until everyone involved in an investigation can agree that it provides a reasonable representation of the incident. If agreement cannot be reached then investigators must select one version of the diagram for further analysis. This decision to move to subsequent stages of analysis is influenced by the scope of the investigation and by pragmatics. For instance, we could extend Figure 3 to consider the circumstances that led to the PLC update. This could only be done if incident investigators can gain access to the PLC supplier.

Second Stage: Causal Reasoning

The second stage again uses a technique that is common to many investigation methods. The aim is to separate causal factors from contextual information. The analysis starts with the event immediately before the outcome. In this case, we might choose to begin with either 'PLC commands blade motors to restart when circuit breaker is reset, switches still in the 'on' position' or with the supposition that the 'Mechanic goes back to work on blade repair without shutting off motor switches'. Investigators must then ask whether the incident would have occurred if that event had not taken place. If the incident would still have happened then the event cannot be considered as a causal factor. For example, the incident would have been avoided if the PLC had not issued the command to restart the motors. Similarly, we can argue that the incident would have been avoided if the mechanic had not gone back to work on the mill without checking the status of the switches. The analysis proceeds backwards from these events looking at earlier and earlier events in the lead-up to the incident. If the incident would still have happened if an event had not occurred then it cannot be considered as a causal factor. For example, the incident might still have occurred even if the foreman and the electrician had not paused to observe the operation of the circuit breaker. Problems can arise from situations in which an incident occurred because something else did not happen. In this case, we must argue that the incident would have been avoided if that event had occurred. For example, can we be sure that the incident would really have been avoided if the Mechanic had switched off the motors? There may be other ways in which the accident could still have happened even if this event had taken place. These difficulties occur because counterfactual reasoning is non-truth functional. In other words, we must make an argument about what could have happened rather than what actually did take place. It can be difficult to validate such assumptions.

Investigators must then map the causal factors that have been identified from the ECF diagram to failures in the IEC 61508 lifecycle requirements that are illustrated in Table 1. One means of doing this is to identify the conditions that contributed to each causal event in the ECF chart. These conditions typically capture latent issues, including development and operation decisions that create the context for E/E/PES-related mishaps. For instance, the PLC command to restart the blade motors when the circuit breakers were reset was made more likely by the lack of adequate risk assessment during the reprogramming of the PLC. This, in turn, was arguably made more likely by the lack of sufficient training material in the conduct of such risk assessments during the maintenance of PLC's in the process industries. Similarly, the mechanic's failure to shut off the motor power was arguably more likely if they assumed that the PLC would not allow an automatic restart. It might also have been made more likely by the fact that the power settings in the control room could not be observed from the mill. The mechanic may have assumed that the foreman had switched off the power when he left to help the electrician. The mechanic's supposed actions were also probably affected by the foreman's failure to inform him that the power supply was not disconnected before he departed. All of these contributory factors were made more likely by the lack of a formal permission to work scheme of lockout procedures

for ad hoc maintenance such as that performed on the circuit breakers. Table 3 presents some of the results of mapping these causal factors back to violations in the IEC 61508 lifecycle requirements. A justification helps others to understand why investigators identified particular problems in the development or operation of the system.

The analysis of the blade mill incident illustrates a number of important points about the cause analysis of accidents involving E/E/PES. In particular, the technical causes that lead to bugs or inadequate testing often form part of more general failures in the operation, maintenance and regulation of safety-critical systems. This observation is common to all of the E/E/PES related incidents we have analyzed in applications ranging from mineral extraction through maritime command and control to the fluidized catalytic cracking of crude oil (Johnson, 2003a). This observation leads to an important requirement for the future development of our work. We have used IEC 61508 lifecycle requirements to provide a causal taxonomy for E/E/PES related incidents. This was motivated by the commercial uptake of the standard and by the organizational objectives of HSE's Electrical and Control Systems Unit. If another taxonomy were to be used in the future then it would also have to capture the range of technical, organization and managerial causes of these accidents. The case study also reveals certain weaknesses in our application of IEC 61508. We have simply used lifecycle requirements from the standard to provide a causal taxonomy for E/E/PES related incidents. The standard does not explicitly address problems in the regulatory environment; this causes particular problems in our analysis of the blade mill incident given the supposed need for greater regulatory support in risk assessment for PLC reprogramming. Similarly, the standard provides no means of identifying failures that were due to weaknesses in the standard itself. This is a significant omission. Incidents can still occur even if an organization satisfies all of the IEC 61508 lifecycle requirements. We are currently addressing these issues by extending the classification illustrated in Table 1. As mentioned in previous sections, our intention is to develop explicit means of providing feedback about these situations in which development standards fail to ensure the safety of an E/E/PES application.

Third Stage: Generating Recommendations

Investigators can use the causal summary chart illustrated in Table 3 to help identify potential recommendations. Table 4 illustrates one format that can be used to document and justify domain and incident dependent recommendations. Each potential intervention is associated with a priority assessment, with an authority responsible for implementing it and with a potential implementation timescale. The information recorded in these recommendation tables can be used to assist in the monitoring of any accident reporting system. For example, electronic information systems are increasingly being used to identify patterns between causal factors and previous recommendations (Johnson, 2003). If the same set of recommendations continues to be used to address the causal factors of similar incidents then regulators may have to intervene to find more effective remedies. It is also important to identify situations in which recommendations are consistently rejected or inadequately implemented.

Tables 3 and 4 are intended to document the process used to investigate more complex incidents. Co-workers, safety managers and regulators should be able to trace back particular recommendations through the previous stages of any causal analysis to identify the reasons why particular interventions are proposed. For example, recommendation 3 proposes the introduction of a physical interlock that might disable the blade motors when someone is working on the mill. This is based on the observation that operations and maintenance assessments had been less than adequate prior to the incident. In particular, these assessments had failed to predict the impact that the PLC reprogramming would have on the motor restart following a power interruption.

Table 3 - IEC 61508 Causal Summary Chart for Case Study Incident

Causal Event	**Associated Conditions**	**IEC 61508 Lifecycle Classification**	**Justification**	**IEC 61508 Common Requirements Violation**	**Justification**
Supposed: Mechanic goes back to work on blade repair without shutting off motor switches.	Supposition: Mechanic may have known about intended operation of the PLC and assumed that it would not allow restart following circuit breaker trip	**Operation and Maintenance:** 4. operations procedures not assessed.	If the mechanic had assumed that the PLC would prevent any automatic restart of the motors following a circuit breaker trip then he was relying on a safety net for a ènormal maintenance procedureê. Hence those procedures should be reassessed.	**Functional Safety Assessment:** 1. LTA operations and maintenance assessment 2. Modification assessment LTA.	The incident may be symptomatic of other problems in operations and maintenance assessment not just in the mill clearing and repair procedures. Similarly, there may be other problems with the assessment of modifications beyond the PLC reprogramming. Deeper questions may have to be raised about the procedures and techniques used to assess functional safety across the plant.
	Supposition: blade motor control settings could not be observed at the mill hence mechanic may have assumed foreman has shut down the mill as before	**Installation and maintenance:** 4. installation design	The layout of the motor controls in the control room prevented the mechanic from easily checking that the foreman had switched them off before leaving to help the electrician. Warning lights could have been located close to the mill to indicate the status of the motor switches.		
	Foreman fails to alert mechanic that mill power supply is not disconnected while they work on the circuit breaker.	**Operation and Maintenance:** 2. permit/hand over procedures need improvement 3. maintenance procedures not impact assessed.	If handover procedures had been in place then the foreman might have informed the mechanic about his intentions on leaving to help the electrician. This should have explicitly addressed the implications of the work on the circuit breaker and on shut-down procedures during any further mill repairs.	**Safety Management:** 1. LTA safety culture 2. LTA safety audits	The hand-over procedure between the foreman, mechanic and also the electrician may be symptomatic of deeper problems with safety management in a small to medium sized enterprise. A safety audit should raise awareness of potential hazards and the safety implications of apparently routine maintenance operations.
	No formal permission to work scheme or lockout procedure for ad hoc maintenance activities.				
PLC commands blade motors to restart when circuit breaker reset and switches still in the ènonê position	Supposition: Need more risk assessment training material for PLC reprogramming in process industries.	**Modification:** 2. LTA manufacturers information. 4. LTA verification and validation.	The company responsible for the PLC update arguably did not appreciate the need to formally consider the implications of the changes on the operation of the mill. Hence the potential restart hazard was not adequately tested for.	**Safety Management:** 3. LTA management of suppliers **Documentation:** 2. documentation incomplete	The reprogramming of the PLC does not seem to have been supported by a detailed consequence assessment. Again, additional documentation may be required from regulatory organisations to guide E/E/PES suppliers about the best means of performing such a hazard assessment. The operators of the mill might also use such guidance to validate any maintenance activities by suppliers.
	Inadequate risk assessment allows PLC reprogramming of restart hazard following power resumption	**Haxzard and Risk assessment:** 1. Consequence and likelihood estimation **Modification:** 1. impact analysis incorrect			

Table 4 - Recommendation Summary Form (LTA – Less Than Adequate)

Causal Event	**Associated Conditions**	**IEC 61508 Lifecycle Classification**	**IEC 61508 Common Requirements Violation**	**Recommendation**	**Priority**	**Responsible authority**	**Deadline for response**	**Date Accepte Rejecte**
Supposed: Mechanic goes back to work on blade repair without shutting off motor switches.	Supposition: Mechanic may have known about intended operation of the PLC and assumed that it would not allow restart following circuit breaker trip	**Operation and Maintenance:** 4. operations procedures not assessed.	**Functional Safety Assessment:** 1. LTA operations and maintenance assessment 2. Modification assessment LTA.	1. Review operations and maintenance procedures to avoid routine reliance on safety net features.	High	Plant safety manager	1/4/2003	Accepted 15/2/2003
	Supposition: blade motor control settings could not be observed at the mill hence mechanic may have assumed foreman has shut down the mill as before	**Installation and maintenance:** 4. installation design		2. Review design of control room to provide operators with control information on mill and associated plant.	Medium	Plant safety manager	1/6/2003	
				3. Consider adding interlock on mill access platform.	High	Plant safety manager	1/6/2003	
	Foreman fails to alert mechanic that mill power supply is not disconnected while they work on the circuit breaker.	**Operation and Maintenance:** 2. permit/hand over procedures need improvement 3. maintenance procedures not impact assessed.	**Safety Management:** 1. LTA safety culture 2. LTA safety audits	4. Introduce and document a formal permit to work scheme for all repair activities.	High	Plant safety manager	1/4/2003	Accepted 15/2/2003
	No formal permission to work scheme or lockout procedure for ad hoc maintenance activities.			5. Develop handover procedures when repair tasks interrupted	High	Plant safety manager	1/4/2003	
PLC commands blade motors to restart when circuit breaker reset and switches still in the èonê position	Supposition: Need more risk assessment training material for PLC reprogramming in process industries.	**Modification:** 2. LTA manufacturers information. 4. LTA verification and validation.	**Safety Management:** 3. LTA management of suppliers **Documentation:** 2. documentation incomplete	6. Develop training material for E/E/PES suppliers and for operators on necessary hazard identification during PLC reprogramming.	Medium	Industry Regulator	1/9/2003	
	Inadequate risk assessment allows PLC reprogramming of restart hazard following power resumption	**Hazzard and Risk assessment:** 1. Consequence & likelihood estimation **Modification:** 1. impact analysis incorrect		7. Conduct formal hazard identification process to determine if there are any additional threats posed by reprogramming of PLC on this plant and supplierês other installations.	High	PLC Supplier Safety Manager	1/6/2003	Accepted 15/2/2003

Conclusions

As mentioned, the UK Management of Health and Safety at Work Regulations 1999 (HSE, 1999) require every employer to carry out a risk assessment, introduce the necessary preventive and protective measures, and monitor these measures. The associated approved code of practice explains that this monitoring includes an obligation to '*adequately investigating the immediate and underlying causes of incidents and accidents to ensure that remedial action is taken, lessons are learnt and longer term objectives are introduced'*. Unfortunately, there are few recognized techniques that companies might use to analyze E/E/PES related incidents. This paper, therefore, introduces two different approaches for this class of adverse events. The first builds on a simple flowchart that helps investigators identify the causes of a mishap by answering a series of questions. These questions guide the causal analysis to identify underlying problems in the design, development or operation of E/E/PES hardware and software. Each failure identified by the flowchart can be related back to lifecycle requirements within the IEC 61508 standard.

We have also described an extended investigation technique that is appropriate for more complex or more critical incidents. Additional stages are introduced to provide intermediate documentation during the causal analysis. Investigators can use these documents to justify recommendations to their peers, to safety managers and to courts of law. This second approach relies upon reconstructions using a simplified form of the US Department of Energy's Events and Causal Factors (ECF) charting. The resulting ECF diagrams help to distinguish contextual information from causal factors. Each causal factors is then analyzed to identify potential failures in the IEC 61508 lifecycle. This is done using a checklist, illustrated in Table 1.

Our use of IEC 61508 is justified because it provides a means of feeding the insights derived from any incident investigation back into the future maintenance and development of hardware and software within safety-critical applications. Our techniques are likely to identify incidents that cannot easily be attributed to lifecycle phases or common requirements in IEC 61508. The link between constructive design standards and analytical investigation techniques can, therefore, yield insights into the limitations of these standards. An implicit motivation in our work is to provide the feedback mechanisms that are necessary to improve the application of standards, such as IEC 61508 and DO-178B. HSE aim to incorporate this work in published guidance material.

Acknowledgements

Thanks are due to Bill Black (Black Safe Consulting) and Peter Bishop (Adelard) for providing comments on the initial draft of this document.

References

Department of Energy, DOE Guideline Root Cause Analysis Guidance Document, Office of Nuclear Energy and Office of Nuclear Safety Policy and Standards, U.S. Department of Energy, Washington DC, USA, DOE-NE-STD-1004-92, http://tis.eh.doe.gov/techstds/standard/nst1004/nst1004.pdf, 1992.

HSE (1999), *Management of health and safety at work* Approved Code of Practice L21 HSE Books.

HSE (2003), P.G. Bishop, R.E. Bloomfield, L.O. Emmet, C.W. Johnson, W. Black, V. Hamilton and F. Koorneef, Learning from incidents: Outline scheme for E/E/PE safety-related systems,

HSE Contract Research Report, Available at http://www.hse.gov.uk/research/crr_htm/index.htm

International Electrotechnical Commission (2000), IEC 61508 Functional Safety of Electrical/Electronic/ Programmable Electronic Safety-Related Systems.

International Electrotechnical Commission (2003), *IEC Functional Safety Zone*, http://www.iec.ch/functionalsafety.

C.W. Johnson (2003 in press), A Handbook for the Reporting of Incidents and Accidents, Springer Verlag, London, UK.

C.W. Johnson (2003a), Incident Reporting and Analysis for Electrical, Electronic and Programmable Electronic Systems (E/E/PES) under IEC61508. See http://www.dcs.gla.ac.uk/~johnson/hse

C.W. Johnson, G. Le Galo and M. Blaize, (2000), Guidelines for the Development of Occurrence Reporting Systems in European Air Traffic Control, European Organisation for Air Traffic Control (EUROCONTROL), Brussels, Belgium.

P. Ladkin and K. Loer (1998), Why-Because Analysis: Formal Reasoning About Incidents, Bielefeld, Germany, Document RVS-Bk-98-01, Technischen Fakultat der Universitat Bielefeld, Germany.

N. Leveson, (2002), A Systems Model of Accidents. In J.H. Wiggins and S. Thomason (eds) Proceedings of the 20th International System Safety Conference, 476-486, International Systems Safety Society, Unionville, USA.

ATTEST: an Automated-Test-Tool Evaluation and Selection Technology

Mr Daniel Rowley; Monash University; Clayton, Victoria, Australia
Dr Sita Ramakrishnan; Monash University; Clayton, Victoria, Australia

Keywords: Forensic Software Engineering, Technology Evaluation and Selection

Abstract

A significant part of software testing process improvement effort pertains to defect prevention, software testing technology change management and software testing process change management. ATTEST is an automated-test-tool evaluation and selection technology developed by the School of Computer Science & Software Engineering (CSSE) at Monash University in Australia to help SMEs (small- to medium-sized enterprises) improve their management of software testing technology change. Although ATTEST has software-process-improvement-oriented application, it can also be used to help forensic software engineers more easily identify candidate equipment for software-intensive incident and accident investigations. The problem with traditional automated-test-tool (or more generally, computer-aided software engineering (CASE) tools) evaluation and selection techniques is that they provide limited visibility/measurement into the selection (acquisition and/or equipping) of automated-test-tools (or CASE tools). In forensic investigations of software-intensive accidents and incidents, it is important that forensic software engineers correctly identify, measure, and collect the data needed to draw valid conclusions regarding technology adoption. Without an automated-test-tool evaluation and selection process that supports completeness and consistency between evaluations and selections, it becomes difficult for forensic software engineers to justify and evidence their software testing technology change management decisions. While most applications of ATTEST are oriented toward the prevention of software failures (or software-intensive incidents and accidents), we aim to demonstrate that ATTEST also has response-orientated application.

Background

The ATTEST: an Automated-Test-Tool Evaluation and Selection Technology project has been partly funded by the School of CSSE at Monash University to assist with software testing technology adoption in organisations (especially focusing on small- to medium-sized enterprises (SMEs)). This project is an add-on project to the TestIT project funded by the Department of Communication, IT & Arts (DCITA) (URL: http://www.dcita.gov.au/Article/0,,0_1-2_1-4_16089,00.html). One of the aims of that project [12] has been to set up a facility for an independent validation and conformance process for existing commercial software testing tools to address industry's current concerns as articulated by small to medium enterprises (SMEs) and software accreditation bodies such as NATA (National Association of Testing Authorities), Australia. An overview of our work is available at http://honeyant.csse.monash.edu.au/index.html.

The Taxonomy of Computer-Aided Forensic Software Engineering

Forensic software engineering is the utilisation and application of software engineering principles, knowledge, expertise and experience for the purposes of the law (negotiation and mediation) or other dispute resolution processes. In particular, forensic software engineering focuses on research and investigation to determine the relevant data and facts following a software-intensive incident or accident. Rowley and Ramakrishnan [2] argue that forensic software engineering is much like independent verification and validation (IV&V). The IEEE

Standard Glossary of Software Engineering Terminology [8] defines independent verification and validation (IV&V) as "verification and validation performed by an organisation that is technically, managerially, and financially independent of the development organisation". The IEEE Standard Glossary of Software Engineering Terminology also defines verification and validation (V&V) as the "process of determining whether the requirements for a system or component are complete and correct, the products of each development phase fulfil the requirements or conditions imposed by the previous phase, and the final system or component complies with specified requirements". According to Rowley and Ramakrishnan [2], forensic software engineering is an outcome of not considering (or mitigating) the likelihood and consequences of software failure whereas IV&V is an upshot of considering (or mitigating) the likelihood and consequences of software failure. Moreover, forensic software engineering involves strict financial, managerial, and technical independence from both clients and suppliers of software-intensive systems. However, despite these differences, forensic software engineering and IV&V both require software testing technologies (techniques and tools) to investigate and report on the correctness of software-intensive systems.

According to Van Wyk and Forno [1], forensic evidence collection processes must be comprehensive, objective, and precise. Automated-test-tools have long been recognised as an effective way to not only improve software development variables such as productivity and product quality but also address the essential difficulty of forensic software analysis: gathering evidence of software failures and faults through clouds of complexity, conformity, changeability, and invisibility. According to Brooks [10], the most difficult work of software engineering is not coding or testing but the essential parts of software engineering. Brook argues that software development is difficult because of the essential complexity, conformity, changeability, and invisibility of software-intensive systems. Moreover, Bruckhaus et al. [3] argue that tools can help improve development processes by facilitating activities that were not practiced before or by supporting activities that are usually carried out with little or no tool support. Schach [6] argues that the simplest form of CASE (computer-aided (or -assisted) software engineering) is the software tool, a product that assists in just one aspect of software production. According to Schach, CASE tools that help the developer during the earlier phases of the process are sometimes termed upperCASE or front-end tools, whereas those that assist with implementation, integration, and maintenance are termed lowerCASE or back-end tools. A CASE workbench is a collection of tools that together support one or two activities, where an activity is a related collection of tasks. Unlike the workbench, which supports one or two activities, an environment supports, at the very least, a large portion of a software process [9].

Our approach to CASE tool evaluation and selection focuses on the V-model which demonstrates how testing activities are related to analysis and design. According to Moriguchi [7], the V-model of software development (see Figure 1.1) is the result of a re-examination of the life cycle model from the point of view of quality assurance. Moriguchi describes the design processes of the V-model as "conversion processes that define [a software solution] in more detail, finally reaching a level of detail that the computer can execute as computer program instructions". We argue that the V-model provides an appropriate process framework for software-intensive incident or accident investigation. Because the V-model details interrelationships between testing and design activities, it is practical for measuring whether or not a developer undertook all reasonable steps to assure software correctness (or more generally, software quality). Furthermore, the V-model can generally be applied to any software development lifecycle and fits into international standard requirements such as ISO 9000.

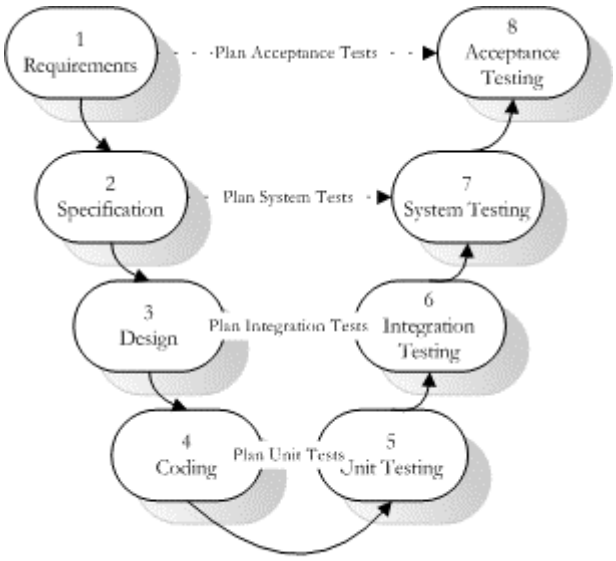


Figure 1.1 - V-Model

By traversing the V-model in a reverse direction (see Figure 1.2) a forensic software engineer is able to appraise the "all reasonable steps"-ness of software product and process documentation. Process documentation that demonstrates "all reasonable steps" to assure software quality is the best defence in software-intensive litigations [14, 15, 16]. Because the V-model mandates that test planning be a part of requirements, specification, design and coding effort, acceptance, system, integration and unit test plans are expected to be compliant with design. In many legal cases, the job of a forensic software engineer is to determine whether or not a software engineer (or team of software engineers) undertook all reasonable steps to ensure that the delivered software product complied with quality requirements. While it could be argued that measuring the compliance of test documentation at different levels of design is tedious work, it is the only way to be sure beyond a reasonable doubt that the developers were or were not negligent with quality assurance (or more specifically, test design and execution). Moreover, if it is not an issue of whether or not the developers were negligent but only whether or not (and how) the software fails, extensive testing still needs to be undertaken to determine the conditions which can cause and caused the software product to miscarry.

While it is obvious that it is difficult (or impossible) to test software fully, automated test tools can help ensure that much of the guesswork/uncertainty and human error is reduced. Furthermore, by viewing or appraising development documentation in the context of the V-model, forensic software engineers are able to better plan and execute their work so that relevant evidence is not excluded. Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless preparation of cumulative evidence. Although using the V-model as a guide for forensic software engineering appears to be a methodical (or breadth-first) heavy-weight analytical approach, it is easy to see that it also accommodates a light-weight inquisitive (or depth-first) style of investigation. That is to say, the tactic does not necessitate that all test results be derived during failure analysis. In some cases, it may be appropriate (or timely) to abandon comb-like search operations (or testing) to concentrate on the validation of a particular hypothesis or casual theory.

Figure 1.3(a) represents a CASE tool that assists with part of the requirements phase (acceptance test planning). Figure 1.3(b) represents a workbench of tools that assist with acceptance test, system test, and integration test planning whereas Figure 1.3(c) depicts an environment that supports all aspects of all phases of the V-model (test planning and test execution). We argue that the forensic software engineering process involves four distinct testing activities: acceptance testing, system testing, integration testing, and unit testing, and four distinct review activities: code review, design review, specification review, and requirements review. When forensic software verification process activities indicate discrepancies between design and test results (poor test coverage), the arrested test specifications need to be corrected and executed or new test specifications need to be designed, written, tested, and executed - in other words, the forensic software validation process begins. In general, the order in which testing activities and review

activities are performed is dependent upon the quantity and quantity of process and product documentation that is made available to the forensic investigator. Nevertheless, we argue that forensic software engineering investigation lifecycles are typified by eight distinct activities. Figure 1.4(a) represents a CASE tool that assists with part of the requirements review phase. Figure 1.4(b) represents a workbench of tools that assist with requirements review, specification, and design review whereas Figure 1.4(c) depicts an environment that supports all aspects of all phases of the forensic V-model (design review and test execution).

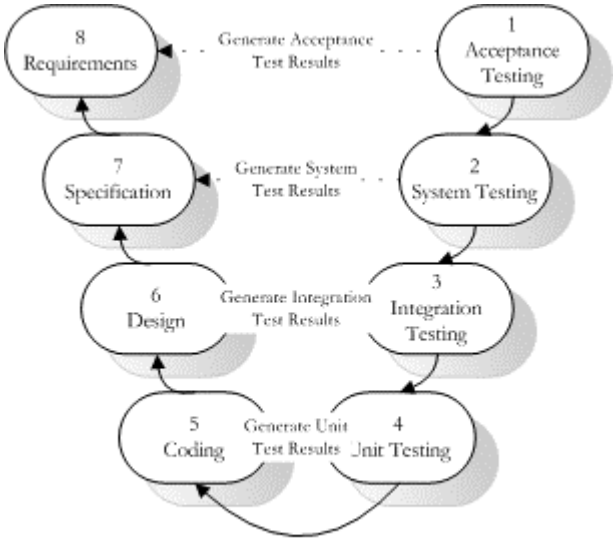


Figure 1.2 - Reverse V-Model

Traditional scorecard systems are traditionally and typically paper-form-based and rely heavily on human effort to construct, validate, maintain and analyse. Furthermore, use of a paper-based system makes it difficult to justify the evaluation and selection of software testing tools when the authenticity of forensic evidence (software failures and faults) is questioned or scrutinised. ATTEST facilitates the mapping of automated-test-tool requirements to automated-test-tool characteristics using a mixture of scorecard evaluation techniques: the evaluation-scorecard technique [13] and the preferred-scorecard technique [13]. The evaluation-scorecard technique and the preferred-scorecard technique have proven to be useful for evaluating and equipping (already-acquired) automated-test-tools however they provide comparatively-minimal insight into whether or not an automated-test-tool acquisition is optimal.

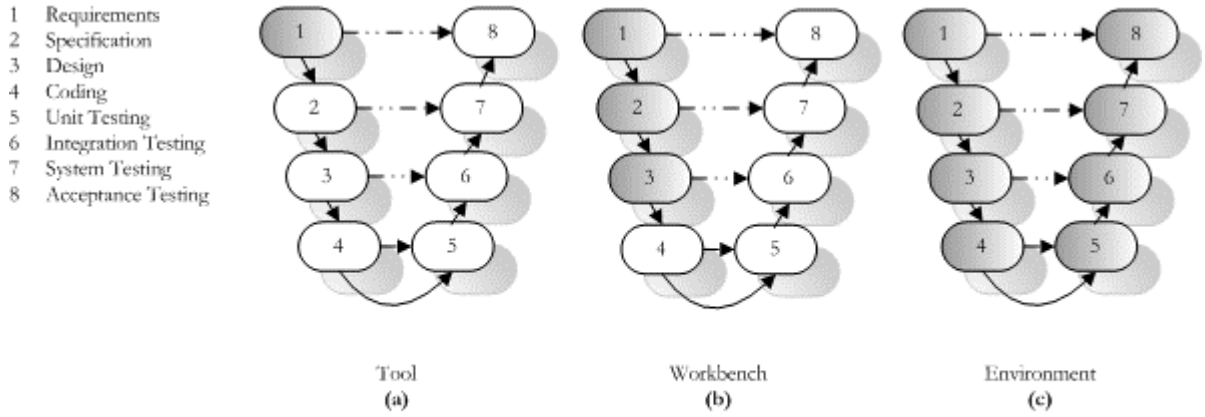


Figure 1.3 - The V-Model and The Taxonomy of CASE

The evaluation scorecard technique involves specifying weighted requirements for an automated-test-tool selection against all characteristics of a technology. On the other hand, the preferred scorecard differs from the evaluation scorecard technique by considering only high-weight requirements (or in other words, highly-preferred characteristics). Both techniques rank

scorecards by their sum of weight-score products however comparative results between both techniques indicate that, in some cases, a CASE tool can obtain two very different rankings using both techniques and the same evaluation scores. The problem with the evaluation-scorecard technique is that it considers all non-zero-weighted characteristics to be essential requirements. In some cases, a CASE tool covering a great number of low-weight requirements can attain a higher ranking than a CASE tool that covers a smaller number of high-weight requirements.

As mentioned earlier, the preferred-scorecard technique only considers high-weight requirements. Another disadvantage of both techniques is that they do not (visibly) separate the specification of requirements from the evaluation of a technology. Moreover, neither technique can distinguish between mandatory (or essential) and optional (or favourable but not essential) requirements. The reason why it is beneficial to distinguish between mandatory and optional requirements is that it allows a forensic engineer (or technology change management groups (TCMGs)) to distinguish between two or more automated-test-tools that cover mandatory requirements equally-well. In some cases (such as the acquisition of new forensic equipment), it may be appropriate to have insight into which automated-test-tools offer additional functionality above that required. In other words, in some situations, it may be appropriate to be cautious about which automated-test-tool characteristics could be required at later dates. On the other hand, in some cases (such as the identification of which already-acquired automated-test-tool to equip), it may not be necessary to separate automated-test-tools that offer additional (but unneeded) features from those that do not.

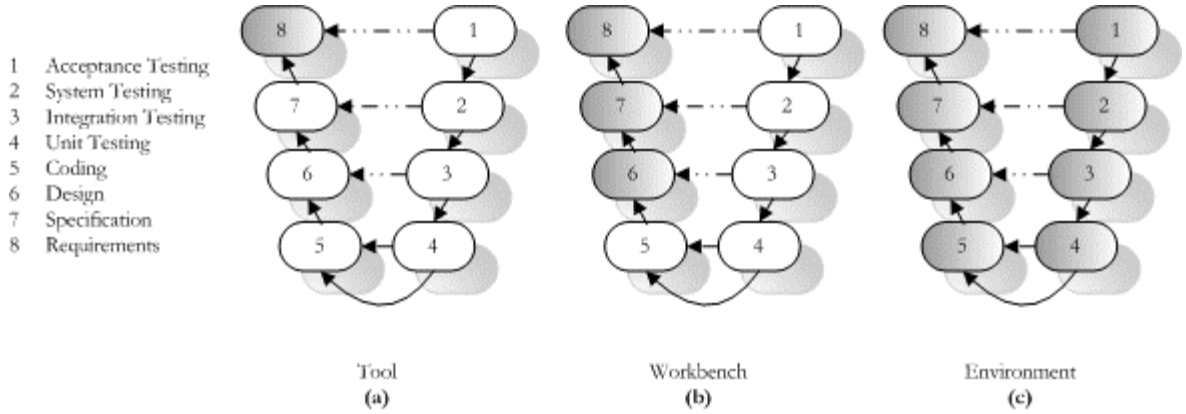


Figure 1.4 - The Reversed V-Model and The Taxonomy of CAFSE

In light of the advantages and disadvantages of the evaluation-scorecard and preferred-scorecard techniques, we propose a new scorecard-technique that allows forensic software engineers (or TCMGs) to specify whether or not a CASE tool requirement is mandatory or optional. Furthermore, our new technique enforces the separation of specification of requirements from the evaluation of a technology; doing so enables requirements specifications and technology evaluations to be reused. Although the evaluation-scorecard and preferred-scorecard techniques both offer a measure of requirement-coverage quality (sum of weight-score products), neither technique explicitly offer a measure of the quantity of requirement-coverage (the percentage of non-zero-weight requirements with non-zero scores). While a trivial computation, a requirements-coverage metric allows forensic software engineers to clearly identify automated-test-tool candidates that satisfy all requirements regardless of rating. In some situations, it may be necessary to select an automated-test-tool based on the quantity and not the quality of the

requirements coverage. Alternatively, it may be appropriate to select an automated-test-tool based not only on the quantity but also the quality of the requirements coverage.

ATTEST: an Automated-Test-Tool Evaluation and Selection Technology

ATTEST is an object-oriented software tool designed to facilitate the evaluation and selection of automated-test-tools that can assist in validating and verifying software products at different levels of design. Brown and Wallnau [4] argue that much of the informality in interpreting any evaluation's results is due to the absence of well-defined goals before starting the evaluation; controlled rigorous techniques for data gathering during the evaluation; and a conceptual framework for analysing the resultant data in the context of existing technologies. In consider of this, we identified two ways to improve the formality in interpreting CASE tool evaluations: by improving the specification of requirements (or definition of goals) and by improving the control and rigor of evaluation data collection. In regard to the controlled, rigorous collection of evaluation data, it is impossible for us (through ATTEST) to provide guidelines on measurement for every feature of all technologies. Instead, we are able to ensure (through design) that evaluators are presented with all the criteria to assess an automated-test-tool against. ATTEST supports the specification of requirements by presenting selectors with all the criteria (or characteristic or features) that can be expected from a particular type of automated-test-tool.

According to Freedman [5], an entity relationship model is a data model that describes attributes of database entities and the relationships among them. Figure 2 depicts an entity relationship model that describes the relational-database entities and entity relationships needed for product-oriented evaluation of CASE tools (products that assists in just one aspect of the production of software). As shown in Figure 1, ATTEST operates on three types of document (or data set): specifications, scorecards, and scoreboards. A specification describes a type of technology (using characteristics) or a set of requirements (using requirements) whereas a scorecard describes the quality of a technology implementation (using scored characteristics). A scoreboard provides sets of measurements for a set of technology implementations (technology scorecards). ATTEST uses folders to collate related specifications, scorecards, and scoreboards. ATTEST can be set up to contain a folder for each distinct software engineering process, activity or task in any software development lifecycle. A technology specification defines the characteristics of a software testing technology (or automated-test-tool) type. A technology specification is a collection of characteristics where each characteristic has a name and a description. On the other hand, a technology scorecard defines the quality of an automated-test-tool in terms of characteristics; in other words, a technology scorecard is an evaluation of an automated-test-tool. A technology scorecard is a collection of scored characteristics where each scored characteristic is a characteristic with a score (between 0 and 100); a characteristic with a high score is a high-quality characteristic whereas a characteristic with a low score is a low-quality characteristic. A technology specification is used to provide evaluators with criteria to assess an automated-test-tool against whereas a technology scorecard is used to enter the results of an automated-test-tool evaluation. A technology scoreboard is a table that lists characteristic coverage metrics pertaining to automated-test-tools: rating, percentage-of-characteristics-covered, and percentage-of-characteristics-not-covered. A rating is a metric that quantifies the quality of an automated-test-tool in terms of characteristic coverage. A percentage-of-characteristics-covered metric describes the quantity of characteristics covered by an automated-test-tool whereas a percentage-of-characteristics-not-covered metric describes the quantity of characteristics not covered by an automated-test-tool. A technology scoreboard is useful for providing evaluators with an overview of the quality of automated-test-tools in a set of automated-test-tools.

While technology scoreboards can provide some insight into which automated-test-tools offer high-quality functionality (characteristic coverage) and/or a high-quantity of functionality, a requirements scoreboard enables forensic engineers to identify those automated-test-tools that provide high-quality coverage of requisite functionality. In some situations, a forensic software engineer may only require a subset of automated-test-tool functionality. In light of this, ATTEST also operates on requirements specifications. A requirements specification describes the requirements of a particular automated-test-tool type. A requirements specification is a collection of requirements where each requirement is a characteristic with a weight (between 1 and 100) and a flag. The weight quantifies the relative importance of the requirement to other requirements whereas the flag indicates whether or not the requirement is mandatory or favourable (optional). A requirement with a weight of zero is not considered to be a requirement (regardless of the typing). The problem with technology scoreboards is that they provide no insight into which automated-test-tools best cover any subset of characteristics. By specifying whether or not a requirement is requisite (mandatory) or favourable (optional), two subsets of requirements can be identified: mandatory requirements and optional requirements. In general, for each subset of requirements, three metrics can be computed (on each automated-test-tool (technology scorecard)): rating, percentage-of-requirements-covered, and percentage-of-requirements-not-covered.

A rating is a metric that quantifies the quality of an automated-test-tool in terms of requirements coverage. A percentage-of-requirements-covered metric describes the quantity of requirements satisfied by an automated-test-tool whereas a percentage-of-requirements-not-covered metric describes the quantity of requirements not satisfied by an automated-test-tool. ATTEST offers two sets of metrics: one set for each subset of requirements (or type of requirement). The set of metrics for mandatory requirements provide insight into which automated-test-tools offer a high-quality and/or high-quantity coverage of requisite automated-test-tool characteristics. On the other hand, the set of metrics for favourable (optional) requirements provide insight into which automated-test-tools offer high-quality and/or high-quantity coverage of favourable (optional) automated-test-tool characteristics. Alike a technology scoreboard, a requirements scoreboard is a table that lists automated-test-tools according to a number of metrics: rating, mandatory rating, optional rating, percentage-of-requirements-covered, percentage-of-requirements-not-covered, percentage-of-mandatory-requirements-covered, percentage-of-mandatory-requirements-not-covered, percentage-of-optional-requirements-covered, and percentage-of-optional-requirements-not-covered.

The benefit of producing a complete specification of a technology (automated-test-tool type) is that it helps forensic software engineers ensure that their specification of requirements is complete and consistent. Although computerisation cannot validate the weighting or typing of automated-test-tool requirements, ATTEST can help ensure that forensic engineers only qualify true automated-test-tool characteristics. The benefit of producing a complete evaluation of an automated-test-tool is that prevents the need for re-evaluation at a later stage (where time may be limited). In regard to requirements specification, ATTEST ensures that selectors are presented with all the criteria (or characteristic or features) that can be expected from a particular type of automated-test-tool.

In terms of maintainability, ATTEST is able to accommodate changes to the specifications of technologies. In some cases, a technology specification may contain an erroneous characteristic (or characteristics) or may omit a characteristic (or characteristics). More significantly, a technology type may evolve over time. At present, ATTEST cascades all additions of characteristics to and updates and deletions of characteristics from technology specifications to technology scorecards. However, modifications to a technology specification often cause one or

more technology scorecards to become outdated (and needy of attention). As the design and development of ATTEST continues, we aim to remain focused on ensuring that the technology solves this and other problems faced by technology evolution.

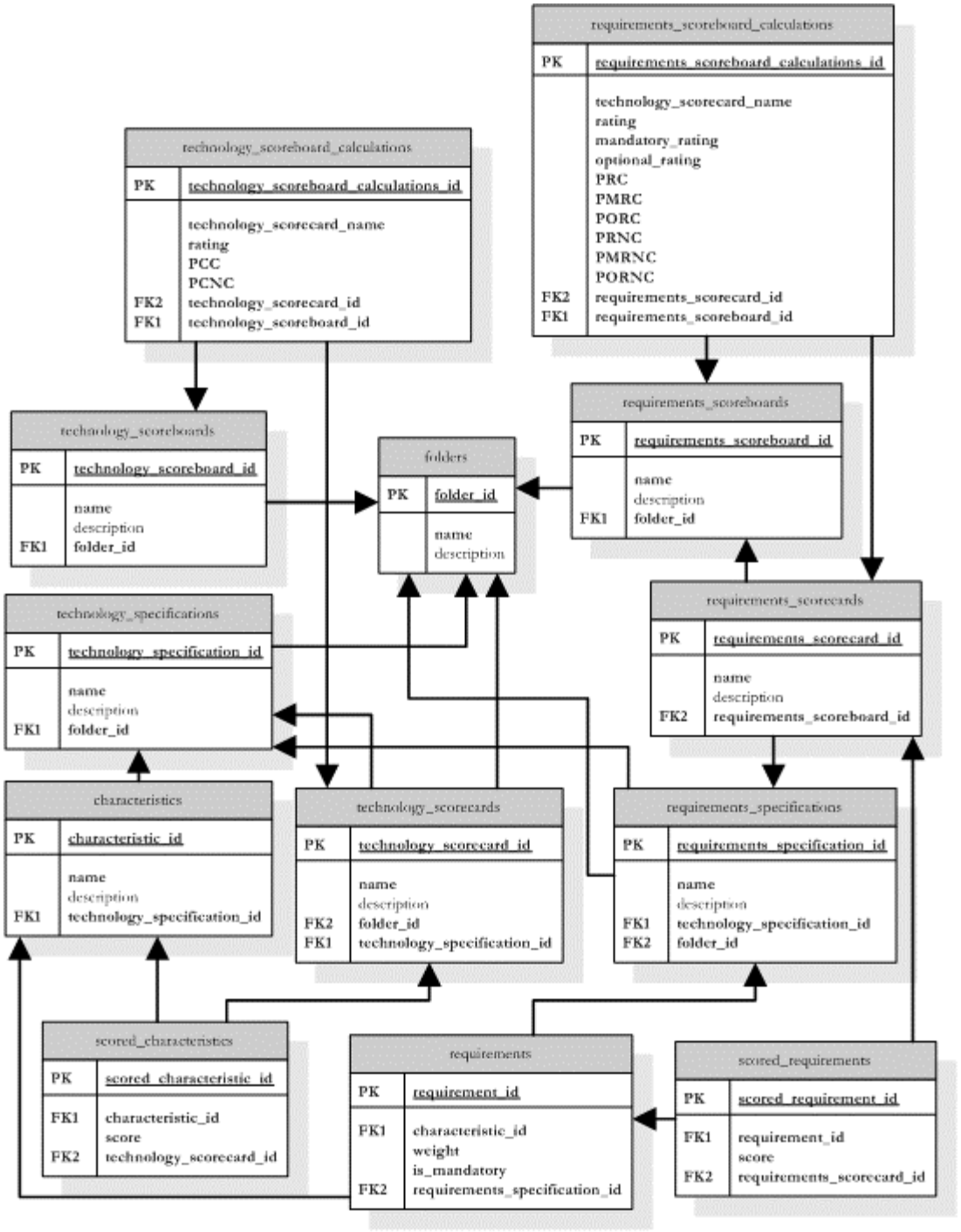


Figure 2 Entity Relationship Model for Computerising the Product-oriented Evaluation of CASE Tools.

Because ATTEST was designed to be general enough to accommodate any software development or software analysis lifecycle model, it is not appropriate to discuss how different CASE tools

support different software development (or forensic software analysis) activity. In fact, many text books describe how CASE tools fit into different parts of the software development lifecycle. Another project within the School of CSSE has elucidated how automated-test-tools support activity in the V-model of software development. Later this year we intend to release a beta version of ATTEST that includes a set of technology specifications that can be used to evaluate automated-test-tools in two contexts: software testing process improvement and software failure investigation and reporting. While it is possible to distribute evaluations of automated-test-tools (given the input/output architecture of ATTEST), we face two problems: not only does the nature of most evaluations tend to be subjective rather than objective (and therefore difficult to validate), some automated-test-tool vendors prohibit the evaluation of their products through stipulations in usage agreements. Nevertheless, forensic software engineering laboratories should perform their own evaluation of equipment as it would help ensure that their selection decisions reflect their understanding of the capabilities and performances of their automated-test-tools.

To realise the function of ATTEST, consider the evaluation of three automated-test-tools (Tool A, Tool B, and Tool C) of type Y that support an activity X (see Table 1.1).

TABLE 1.1 – TECHNOLOGY SCORECARD

Characteristic	Tool A	Tool B	Tool C
characteristic 1	98	0	100
characteristic 2	34	100	87
characteristic 3	56	50	12
characteristic 4	0	97	78

Each evaluation of a tool scores four characteristics that describe Y. While it is relatively trivial to calculate the rating of and the percentage of characteristics covered by each tool in this example (see Table 1.2 and Figure 3.1), it becomes much more difficult when there are tens or hundreds of characteristics to consider.

TABLE 1.2 – TECHNOLOGY SCOREBOARD

	Tool A	Tool B	Tool C
Rating	(98 + 34 + 56 + 0)/400 = 0.4700	(0 + 100 + 50 + 97)/400 = 0.6175	(100 + 87 + 12 + 78)/400 = 0.6925
Percentage of characteristics covered	¾ = 0.7500	¾ = 0.7500	4/4 = 1.000
Percentage of characteristics not covered	¼ = 0.2500	¼ = 0.2500	0/4 = 0.000

In fact, it takes O (n)-time to compute the rating and percentage of characteristics covered by a tool. Moreover, consider the requirements of a tool of type Y to support an activity X (see Table 1.3 and Figure 3.2). Table 1.3 specifies three requirements of a tool of type Y to support an activity X: two heavy-weight mandatory requirements and one mid-weight optional requirement.

TABLE 1.3 – REQUIREMENTS SPECIFICATION

Characteristic	Requirement Weight	Requirement Is Mandatory
characteristic 1	100	true
characteristic 2	60	false
characteristic 3	90	true
characteristic 4	0	false

Calculating the ratings of and the percentages of characteristics covered by each tool in this example (see Table 1.4) is somewhat laborious given the complexity of the computations. Although the time complexity to compute ratings of automated-test-tools in a requirements context is also proportional to the number of characteristics (mapped to requirements), the calculations are much more intricate. Again, it is easy to see that the measurement process becomes more laborious as the number of characteristics to consider increases. In other words, without computerisation, measuring the suitability of a CASE tool selection requires substantial routine (and error-prone) effort.

TABLE 1.4 – REQUIREMENTS SCOREBOARD

	Tool A	Tool B	Tool C
Rating	(100 * 98 + 60 * 34 + 90* 56) / (100 * 100 + 60 * 100 + 90 * 100) = (9800 + 2040 + 5040) / (10000 + 6000 + 9000) = 16880 / 25000 = 0.6752	(100 * 0 + 60 * 100 + 90 * 50) / (100 * 100 + 60 * 100 + 90 * 100) = (0 + 6000 + 4500) / (10000 + 6000 + 9000) = 10500 / 25000 = 0.4200	(100 * 100 + 60 * 87 + 90 * 12) / (100 * 100 + 60 * 100 + 90 * 100) = (10000 + 5220 + 1080) / (10000 + 6000 + 9000) = 16300 / 25000 = 0.6520
Mandatory rating	(100 * 98 + 90 * 56) / (100 * 100 + 90 * 100) = (9800 + 5040) / (10000 + 9000) = 14840 / 19000 = 0.7811	(100 * 0 + 90 * 50) / (100 * 100 + 90 * 100) = (4500) / (10000 + 9000) = 4500 / 19000 = 0.2368	(100 * 0 + 90 * 50) / (100 * 100 + 90 * 100) = (4500) / (10000 + 9000) = 4500 / 19000 = 0.2368
Optional rating	(60 * 34) / (60 * 100) = 2040 / 6000 = 0.3400	(60 * 100) / (60 * 100) = 6000 / 6000 = 1.0000	(60 * 87) / (60 * 100) = 5220 / 6000 = 0.8700
Percentage of requirements covered	3/3 = 1.0000	2/3 = 0.6667	3/3 = 1.0000
Percentage of mandatory requirements covered	2/2 = 1.0000	1/2 = 0.5000	2/2 = 1.0000
Percentage of optional requirements covered	1/1 = 1.0000	1/1 = 1.0000	1/1 = 1.0000
Percentage of requirements not covered	0/3 = 0.0000	1/3 = 0.3333	0/3 = 0.0000
Percentage of mandatory requirements not covered	0/2 = 0.0000	1/2 = 0.5000	0/2 = 0.0000
Percentage of optional requirements not covered	0/1 = 0.0000	0/1 = 0.0000	0/1 = 0.0000

Not only does ATTEST expedite the measurement process, it also facilitates the interpretation and presentation of measurement data. While it is not difficult work to replicate (or duplicate) and reorder data, it is a tedious process that is better managed by computer technology. ATTEST operates a SQL (Structured Query Language) interface (to an implementation of the entity relationship model (in Figure 2)) that not only allows a forensic engineer to enter evaluation and selection (requirements) data but also customise the presentation of report data.

TABLE 1.5 – REQUIREMENTS SCOREBOARD

Candidate	Rating	Mandatory rating	Optional rating	%RC	%MRC	%ORC	%RNC	%MRNC	%ORNC
Tool A	0.6752	0.7811	0.3400	1.0000	1.0000	1.0000	0.0000	0.0000	0.0000
Tool B	0.4200	0.2368	1.0000	0.6667	0.5000	1.0000	0.3333	0.5000	0.0000
Tool C	0.6520	0.2368	0.8700	1.0000	1.0000	1.0000	0.0000	0.0000	0.0000

Table 1.5 and Figure 3.3 display the measurement data from Table 1.4 as presented in ATTEST by default. While Table 1.5 contains much useful information, Miller's law [11], states that at any one time, a human being can concentrate on at most 7 ± 2 quanta of information. In light of this, ATTEST can be controlled to display and order any subset of data columns and records (rows). In continuance of our example, Table 1.6 and Figure 3.4 show a result of using stepwise-refinement and SQL to display the pertinent data needed to select a tool that best covers all (100% of) the requirements for a tool of type Y.

TABLE 1.6 – FILTERED REQUIREMENTS SCOREBOARD

Candidate	Rating	Mandatory rating	Optional rating	%RC	%MRC	%ORC
Tool A	0.6752	0.7811	0.3400	1.0000	1.0000	1.0000
Tool C	0.6520	0.2368	0.8700	1.0000	1.0000	1.0000

Conclusions

At present, ATTEST only supports the evaluation and selection of the simplest form of CAFSE (or more generally, CASE): the software tool. That is, ATTEST is known to be better suited to guiding the acquisition and/or equipping of automated-test-tools than to the acquisition or equipping of automated-test-workbenches or -environments. Furthermore, ATTEST only supports the product-oriented evaluation of CASE tools; future directions for ATTEST aim to not only accommodate the product-oriented evaluation of automated-test-workbenches and -environments but also the process-oriented evaluation of automated-test-tools, -environments and -workbenches. Product-oriented evaluation involves selecting among a set of products that provide similar functionality whereas process-oriented evaluation involves assessing the impact of a new technology on existing practices to understand how it will improve performance or increase quality.

In reality, it is difficult to orthogonally classify automated-test-tools because most modern automated-test-tools support more than one part of the software development (or forensic software analysis) lifecycle. Tools that support more than one software engineering process or task can only be accommodated in ATTEST by producing separate specifications of the tool for each distinct supported process or task. Once a forensic software engineer has identified what task (or type of test planning or test execution) needs to be performed, the engineer can use ATTEST to identify the most appropriate automated-test-tool for that particular task. Again, ATTEST (at this stage) cannot manage with the complexity of identifying optimal automated-test-tool sets (or workbenches or environments) for performing multiple distinct forensic software engineering tasks. Although ATTEST has many useful features (including SQL (structured query language) interfaces and data exportation), it is clear that further work is needed to extend ATTEST into a totally-effectual CASE technology evaluation and selection tool.

Brown and Wallnau [4] maintain that software technology selection, application, and introduction requires consideration of initial technology acquisition cost; long-term effect on quality, time to market, and cost of the organisation's products and services, when using the technology; training and support services' impact of introducing the technology; relationship of this technology to the organisation's future technology plans; and response of direct competitor organisations to this new technology. Although non-technical factors such as acquisition cost are important considerations (in general), ATTEST was designed with focus on ranking automated-test-tools according to their satisfaction of technical (or functional) requirements. Although it is easy to specify non-technical requirements in ATTEST by adding non-technical characteristics into tool specifications, care must be taken to ensure that the weights of non-technical requirements are in proportion to the weights of technical requirements. Alternatively, ATTEST is able to persist delimited-textual shortlists of candidate automated-test-tools into plain text files that can be manipulated by spreadsheet and work processing software. More significantly, we aim to extend ATTEST to allow engineers to pipe shortlists of candidate automated-test-tools back into the short listing process with new requirements specifications so that requirements (or sets of requirements) can not only carry weight but also precedence (or an ordering of importance). The scope of the ISO/IEC 14102:1995 (Information Technology - Guidelines for the Evaluation and Selection of CASE Tools) International Standard is to establish processes and activities to be performed when evaluating different CASE (computer-aided software engineering) tools and selecting the most appropriate for a given organisation and/or project. Although ATTEST was derived from an intention to improve the change management of software testing technologies, it is general enough to be adapted to help evaluate and select COTS (Commercial-Off-The-Shelf)

software components and other types of CASE tools; another direction of ATTEST aims to investigate the feasibility of attaining compliance with ISO/IEC 14102:1995 and other technology evaluation and selection standards.

Using a computerised automated-test-tool evaluation and selection system is an important consideration for forensic software engineering laboratories. Computerising the automated-test-tool evaluation and selection process can help improve the investigation and reporting of software-intensive incidents and accidents because it enables forensic software engineers to more completely and consistently specify their equipment requirements. Although the mapping of characteristics to requirements is a trivial concept, it is a central concept in the design of automated-test-tool evaluation and selection systems that must ensure completeness and consistency between evaluations and selections. This paper has presented and attested a database entity relationship model that describes the database entities and entity relationships needed for computerising the product-oriented evaluation of automated-test-tools (or more generally, CASE tools). Through the demonstration of ATTEST, we aim to prove that regardless of orientation (prevention or response), computerising (and making more formal) the product-oriented evaluation of CASE tools can more easily and more quickly provide confidence in automated-test-tool selections.

References

1. K Van Wyk and R Forno: *Incident Response*, First Edition, O'Reilly & Associates, July 2001
2. D Rowley and S Ramakrishnan: *Forensic applications of software analysis*, in S Lesavich (ed), Proceedings of the Third International Conference: Law and Technology, Cambridge, USA, 6-7 November 2002, ACTA Press, Anaheim, USA, ISBN: 0-88986-333-4, pp 129-134
3. T Bruckhaus. et al: *The Impact of Tools on Software Productivity*, IEEE Software, September 1996, pp 29-38
4. A Brown and K Wallnau: *A Framework for Evaluating Software Technology*, IEEE Software, September 1996, pp 39-49
5. A Freedman: *The computer glossary: the complete illustrated dictionary*, Eighth Edition, AMACOM, 1998
6. S Schach: *Object-oriented and Classical Software Engineering*, Fifth Edition, McGraw-Hill, 2002
7. S Moriguchi: *Software Excellence: A Total Quality Management Guide*, Productivity Press, 1997
8. The IEEE Standard Glossary of Software Engineering Terminology
9. A Fuggetta: *A Classification of CASE Technology*, IEEE Computer, Volume 26, December 1993, pp 25-38
10. F Brooks: *No silver bullet: Essence and accidents of software engineering* (Computer, 1987); in F. P. Brooks, Jr.: The Mythical Man-Month: Essays on Software Engineering, Addison-Wesley, Reading, Mass., 1995, pp 177-203
11. G Miller: *The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information*, The Psychological Review, March 1956, pp 81-97
12. S Ramakrishnan and C Mingins: *A Faciltity for Conformity and Compliance Testing*, Commonwealth Government Australia, Department of Communication IT and the Arts (DCITA) two-year funded Test-IT project, June 2001
13. E Dustin, J Rashka, and J Paul: *Automated Software Testing: Introduction, Management, and Performance*, Addison-Wesley, 1999
14. J Cosgrove: *Software Engineering and the Law*, IEEE Software, Volume 18, Number 3, 2001

15. B Lawson: *An Assessment Methodology for Safety Critical Systems*, bud@damek.kth.se
16. T DeMarco and T Lister: *Both Sides Always Lose: Litigation of Software-intensive Contracts*, Crosstalk, Volume 13, Number 2, 2000

JLab Web Based Tracking System for Integrated Incident, Accident, Inspection, and Assessments*

S. Prior; REM, CHMM; Jefferson Lab; Newport News, Virginia, USA
R. Lawrence; Jefferson Lab; Newport News, Virginia, USA

Abstract

The Thomas Jefferson National Accelerator Facility, or JLab, is a Department of Energy particle accelerator used to conduct fundamental physics research. In such a facility there are numerous statutory, regulatory, contractual, and best practice requirements for managing and analyzing environmental health and safety (EH&S) related data.

A tracking system has been developed at JLab that meets the needs of all levels of the organization, from the front line worker to the most senior management. This paper describes the system implementation and performance to date.

Introduction

The ability to accurately track EH&S related findings from first encounter to closure is an important tool in the EH&S professional's arsenal. The DOE requires national lab management contractors to implement numerous programs related to EH&S tracking and reporting. Reports include annual self-assessments, accident and incident reports, operational readiness assessments, EH&S inspection program results, and trending [6]. The lab management requires trend analysis and corrective action tracking to closure on all findings. The EH&S professional requires historical data related to incidents and inspections within a given area or related to a specific individual.

Items are tracked in order to identify trends and provide lessons learned. The intent is to prevent future occurrences of events with similar root causes. Any of the above may also produce information that can be used to develop just in time or topic specific training.

Each kind of report has seemingly disparate requirements depending on the intended audience. In fact, there are report tracking software packages available (at great expense), but these are customized for specific industries or tasks. The authors are not aware of integrated packages that provide the ability to enter and track all of the information needed for a comprehensive EH&S trending analysis.

There were a great many similarities in the basic information required to accurately analyze trends in related functions. For example, work areas or zones within the facility or the management chain for an area or employee will require similar input data, e.g. system owners, personnel supervisors, and location.

Investigation of seemingly unrelated, recent near miss events at JLab uncovered an interesting pattern. Although a strong environmental health and safety (EH&S) program was in place, in practice the program had gaps in two key areas: 1.) Clear lines of communication of EH&S related information were not in place to relay lessons learned, and 2.) Clear lines of management responsibility for hazardous tasks and the personnel that performed them were missing or

* This work was supported by US DOE Contract No. DE-AC05-84-ER40150.

ambiguous. A classic example of the latter is an employee of one department matrixed to another department to perform some type of work. In this scenario, there is a potential for neither the direct supervisor nor task supervisor to take responsibility for performing the task hazard analysis, since each may assume that the other has already completed it.

An initiative was developed in response to these inefficiencies. This initiative would accomplish the following:

- Increase worker awareness of EH&S related action items
- Increase management awareness and ownership of EH&S corrective actions
- Facilitate better communication of EH&S related issues and lessons learned at all levels
- Identify areas where additional focused training may be required

In investigating how these goals may be accomplished, it was recognized that one must be able to view and track a comprehensive set of EH&S related information and have the information integrated and available for interactive research [8]. It is from this initiative the tracking system was developed.

Development

The JLab EH&S tracking system was originally developed as a tool to aid the EH&S Tracking, Trending, & Training (T3) office in the capture and reporting of periodic EH&S inspections. These inspections, along with their associated findings and corrective actions, had originally been entered in a paper logbook that was inaccessible to personnel affected by the finding. The initiative for ensuring a corrective action was closed fell on the line manager for the area. However, the line manager, and indeed the management chain, did not have ready access to the inspection findings. By the time data made its way to senior management in weekly or monthly reports, the information was highly edited. In addition, the inspection reports did not include data that was relevant to similar but a broader class of findings since they were single event driven.

At the same time that the initial inspection tracking system was in the requirements development phase, it was observed that there was a similar strong need for tracking accident and incident investigations and corrective actions. Accident and incidents share many of the same attributes as inspections, but they also have a uniquely significant data subset.

Requirements for the tracking system grew to cover four main functional areas:

- Inspection (includes laser audits)
- Accident/Incident Investigation
- Assessments
- Radiation Deviation Reporting

From these functional areas, a set of requirements was developed that strived to make use of the system as ubiquitous and utilitarian as possible. The requirements are outlined in the following objectives:

- Work across multiple platforms
- Be useable by personnel with limited computer skills
- Viewable by anyone
- Output to other data management tools
- Import data from existing personnel and area databases

- Support trending analysis
- Be secure against unauthorized use
- Support group or individual e-mail notification of affected personnel
- Support electronic logging (e-log) of entries

In addition to the basic functionality, emphasis was placed on the customer requirements. The paper system had been neglected due to lack of accessibility and the T3 system would be no different if it was difficult to access or navigate. Given the cross facility needs for EH&S information, the list of system customers covered a wide spectrum.

Customers:

- Safety wardens
- Professional EH&S Staff
- Line management
- Line management
- Division/department management
- Senior Executive management
- Oversight/Compliance Agencies

Input was solicited from potential users at all levels. In addition, information was gleaned from paper forms in order to present the field worker with familiar form entries. A less obvious requirement was development of the T3 system as inexpensive as possible.

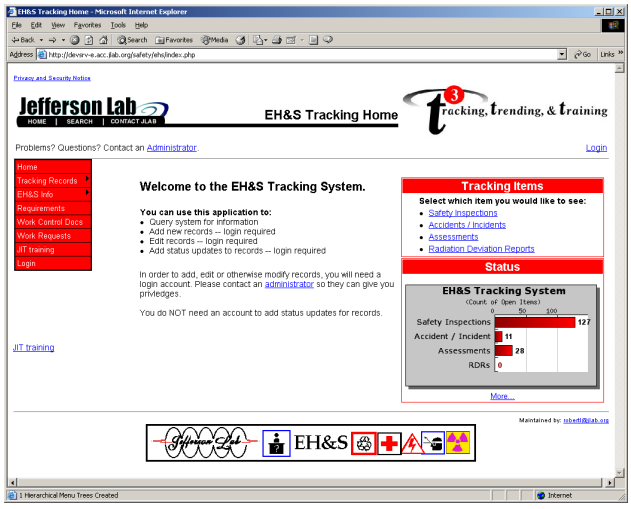


Figure 1 -Tracking System Home Page

System Overview

In order to meet some of the requirements described earlier, it was determined that the system should be implemented in a web-based, client-server environment over the JLab intranet. An overview of the functionality of the T3 tracking system starts with the home page, shown in (figure 1). In addition to normal menu selection items, a set of links related to tracking items is

presented to the user. For quick reference there is also a bar graph of open items (draws the curiosity of the user to find out if they are in the hot seat).

Each page of the T3 system is 'built' by the server at the time the page is requested by the client web browser. Embedded within the HTML code that generates each page are calls to a server-side scripting language known as PHP. The PHP scripts fill in variable data such as pull down menus with the most current data in the system database. This is a powerful tool that allows customization of the information contained in the web interface.

This interface resides on an Apache web server within the lab's intranet. Apache HTTP server and the PHP scripting language are open source server applications from the Apache[1] Software Foundation.

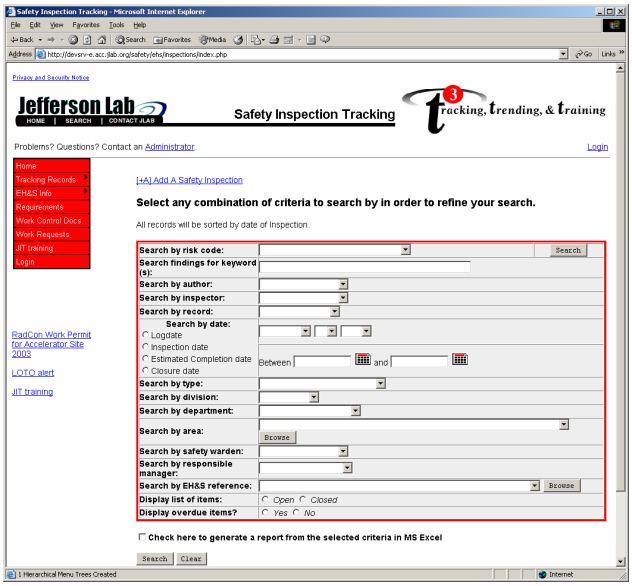


Figure 2 - Safety Inspection Interactive Web

When the web server receives a request for a PHP page, it tells the PHP engine to build it. PHP reads the code in the page and builds a context based HTML page depending on a number of factors; including the type of login, input variables, and if required, executed database queries. This HTML is then given back to the server in order to deliver it to the client's browser.

There are presently four links within the Tracking Items link box. Each one will bring up a context dependent database query dialogue box tailored to the task. For example, an inspection dialogue form (figure 2) may include information on safety wardens that are linked to a specific area. This would not be an appropriate entry for an assessment database query because assessments are not related to individual areas of the lab. Common items, such as the responsible manager, are available among various dialogues. Reports are available as a printable web page or may be opened in an Excel® spreadsheet for further analysis.

[1] Apache Software Foundation http://www.apache.org/

Inspection items include formal, informal, and incidental findings [1, 7]. Each finding includes a risk matrix assessment of the potential harm that could be caused if the item were left unmitigated. Formal findings are generated during the monthly to quarterly scheduled inspections performed by professional EH&S staff. Informal findings are identified by safety wardens.[2] Incidental findings are observations made by EH&S professionals and line managers in the course of performing other work activities.

Accident/Incident items include tracking corrective actions and the lessons learned from these events [2-4]. Prior to implementing the tracking system, it was noted that the closure process tended to break down once the accident report was transmitted to DOE. Since implementing the tracking system, accident rates have improved through timely and effective closure of corrective actions and potential areas for improvement to the accident investigation process have been noted. Moreover, by correlating the root cause data to DOE's Integrated Safety Management System (ISMS) principles and core values within each finding record, gaps in JLab's ISMS program implementation have also been identified and targeted for corrective action. A nice feature that was added to the system is a hyperlink associated with each record that connects the user to an electronic copy of the relevant accident/incident report.

Assessments include findings from two sources: periodic line self-assessments of EH&S performance and independent assessments [9, 10]. Line self-assessments are subjective performance evaluations prepared by department heads. They provide the departments the unique opportunity to identify their accomplishments while at the same time identifying areas for improvement in both operations and EH&S performance. Independent assessments are objective evaluations of a department's EH&S performance. The Office of Assessment that reports to the Lab Director performs them.

Radiation deviation reports (RDR) include data and corrective actions related to deviation from radiation control processes [5]. These deviations fall below the threshold for external reporting.

However, if they are not effectively resolved and repeat findings occur, the RDR could be elevated to an external oversight agency for follow up.

One important requirement of the system is to ensure that ownership of an open item, from introduction to closure, belongs to a single individual. To ensure this, a responsible manager is assigned to each finding to guarantee that there is one point of contact for the closure of every item.

Integration with E-log and E-mail

The JLab scientific operations make substantial use of electronic log books for documentation of machine operations. Several sub logs also exist, which are tied to specific operational areas or pieces of equipment (figure 3). To facilitate communication of EH&S related information, an EH&S sub-log was created. Users may enter log activity independently into the tracking system

[2] Safety wardens are appointed by their supervisors and are drawn from the pool of employees routinely working at a specific work area location. They serve as an extension of the professional EH&S staff in the day-to-day EH&S oversight of their work area and the people accessing it. They can take the initiative to fix the problem, log the problem and see the solution through to closure.

or they can automatically generate an associative e-log entry when a T3 item is created or updated [8].

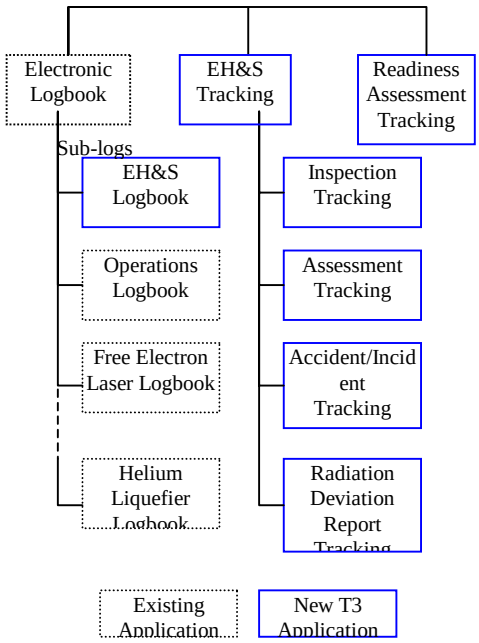


Figure 3 - Hierarchical Diagram of T3

Both the T3 forms and the e-log systems have the ability to e-mail interested parties. E-mail messages are generated automatically, but the user selects the respondents to which they wish to notify about a particular item.

Readiness Assessment

An offshoot of the T3 system was a safety readiness review tracking system. All DOE accelerators are required to perform a safety assessment and identify and track hazards and associated mitigations for any new facility or operation. In the development of such a document, hundreds to thousands of safety related items might be identified and tracked to closure. To ensure that open items are addressed in a timely manner, the tracking system must also track progress towards resolution of safety issues and concerns.

The T3 ReAD (Readiness Assessment Documentation) tracking system does just this using a matrix layout and color coding [11]. The matrix items are interactive with changes and updates added by a mouse click on any underlined item in the matrix. Changes and updates are posted in real time so that the matrix always represents the most current information.

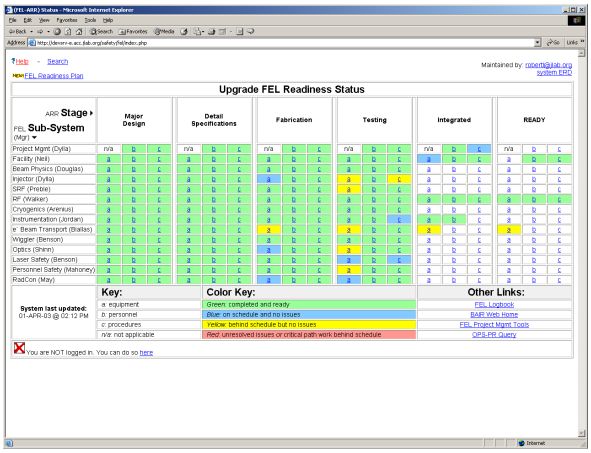


Figure 4 -FEL Readiness Assessment Matrix

Design

A functional diagram of the T3 tracking system is shown in (figure 5). Anyone connecting from within the lab Enterprise intranet may access the Apache web server that delivers the PHP pages for the system. These connections are allowed to pass through onto the Controls intranet because the firewall separating the two networks is configured to allow this. However, the Oracle database server only allows connections from within the Controls intranet, so users must first connect to the Apache web server in order to access the database for the system.

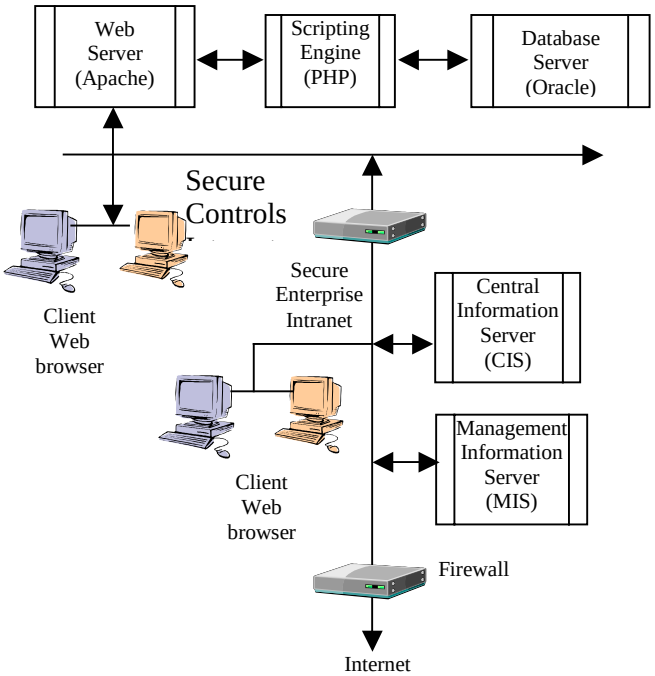


Figure 5 - System Connectivity Diagram

Data Entry and Security

Data can be viewed by anyone within the organization firewall, allowing the system to remain independent from outside the lab, but creation of new entries requires that the user be authorized to use the system. To ensure this authorization, the database keeps a table of authorized users. In order to make a new entry or modify an existing entry, the user must login to the system. Logins are authenticated by a separate and secure section of the web server so that user passwords cannot be sniffed. The header of each page contains a link for the user to log in to the secure server. The user provides his or her username and password over a secure connection to complete this process. If the username and password is not valid, an error message is returned. Once logged in, the user may update or create records for which they are authorized. Administrators are the only people who may delete entries.

Each browser form runs a simple JavaScript function on the client which checks for input completeness before sending the form variables to the server. These variables are used to input the data into the database. This client-side check is done in order to avoid having to tie up the server with these tasks.

For example, when the client opens the inspection form, the PHP engine builds the HTML for the pull down menus dynamically. This information is based on current information contained in the central information server, the management information server, and the Oracle database for the application. These databases contain relational information such as employee/department/division/supervisor/e-mail/inspection type/address and so on. The user then selects the appropriate options from these menus and inputs data into other text boxes. When the user submits this information to the server, the JavaScript checks to see if the required fields are filled out. If they are, the form variables are sent back to the server and their values are added to the database.

Conclusions

A web based tracking system has been developed at Jefferson Lab that meets the needs of a diverse set of customers – from casual to EH&S professional. The system has been in beta test for 6 months and in use for another 6 months. The reception to date has been very enthusiastic and the system is in wide use through out the facility.

References

1. Chapter 5100, Internal Inspections, Jefferson Lab EH&S Manual, April 13, 2000.

2. Chapter 5200, Incident/Injury Investigation, Jefferson Lab EH&S Manual, October 15, 1999.

3. Chapter 5300, Occurrence Reporting, Jefferson Lab EH&S Manual, October 1, 2002.

4. Technical Appendix 5300-T3, Notable Event and Notification Procedure, Jefferson Lab EH&S Manual, October 1, 2002.

5. Technical Appendix 5300-T4, Worker Radiation Protection Rulemaking (10 CFR 835) Reporting, October 1, 2002.

6. Chapter 5400, Documentation and Recordkeeping, Jefferson Lab EH&S Manual, October 1, 2002.

7. Chapter 6410, Laser Safety, Jefferson Lab EH&S Manual, June 24, 1997.

8. Thomas Jefferson National Accelerator Facility Integrated Safety Management System Plan, Rev.6, December 30, 2002.

9. Jefferson Lab Quality Assurance Program Manual, Rev. 4, September 2000.

10. Jefferson Lab Self-Assessment Program Manual, Rev. 6, December 2000.

11. Free Electron Laser Readiness Plan, April 2003.

Biography

Sandra L. Prior, TJNAF; Newport News, Virginia, U.S.A.; telephone +1.757.269.7591; fax +1.757.269.7352; e-mail – prior@jlab.org

Ms. Prior is the Jefferson Lab (JLab) Accelerator Division EH&S principle responsible for tracking, trending, and training. She is also the JLab EH&S Policies and Manual Group Manager. Over her career, Sandra has developed numerous database applications for environmental health and safety related applications. She is the principle designer and author for the database applications described in this paper.

Sandra is a member of the adjunct faculty at Old Dominion University in the Environmental Health Department School of Health Sciences. She is on the Adjunct faculty at the U.S. Army Command & General Staff College (CGSC) in Ft. Leavenworth, KS. She is a member of the Environmental Health Advisory Council for environmental health programs at ODU. She is a member of the United Who's Who Registry of Executives and Professionals. She is also a Registered Environmental Manager and a Certified Hazardous Materials Manager, Masters Level.

Robert L. Lawrence, TJNAF; Newport News, Virginia, U.S.A.; telephone +1.757.269.7331; fax +1.757.269.7352; e-mail – robertl@jlab.org

Mr. Lawrence is a recent Computer and Information Systems program graduate from Christopher Newport University. He has worked two years at Jefferson Lab as a student intern. He is the principle technical author of the EH&S database and associated tools. Robert is proficient in several computer languages, database programming, and scripting languages.

The Creation of an Aviation Safety Reporting Culture in Danish Air Traffic Control

Peter Majgård Nørbjerg; Head of incident investigation, Naviair; Kastrup, Denmark

Keywords: Confidential reporting, non-punitive reporting, Air Traffic Control, reporting culture

Abstract

In 2001, a new law was passed by the Danish Parliament, mandating the establishment of a compulsory, strictly non-punitive, and strictly confidential system for the reporting of aviation incidents. A particular and perhaps unusual feature of this reporting system is that not only are employees (typically Air Traffic Controllers and pilots) ensured strict immunity against penalties and disclosure but also, in fact, any breach against the non-disclosure guarantee is made a punishable offence.

This paper will explore the experience gained during the political process of passing such a law, as well as the practical lessons learned, during the implementation phase of the non-punitive confidential reporting culture in Danish Air Traffic Control.

Introduction

The benefit of flight safety reporting systems to flight safety has been recognised for many years and many systems are in operation today in the North America, Europe, Australasia and elsewhere. Most of these systems share as a common feature that reports are anonymous and aviation personnel who submit reports do so on a voluntary basis. A few systems – such as the ASRS and the CHIRPS makes it possible to report incidents without risking legal action under certain circumstances.

As opposed to these systems, the recently introduced system in Denmark is a *mandatory, non-punitive, and yet strictly confidential* system. The reporting system is mandatory in the sense that air traffic personnel is obliged to submit reports of events, and it is strictly non-punitive in the sense that they ensured indemnity against prosecution or disciplinary actions for any event they have reported.

Furthermore the reporting system is strictly confidential in the sense that the reporter's identity may not be revealed outside the agency dealing with occurrence reports. Reporters of incidents are ensured immunity against any penal and disciplinary measure related to an incident if they submit a report of within 72 hours of its occurrence and if it does not involve an accident or does not involve deliberate sabotage or negligence due substance abuse (e.g., alcohol). Moreover, punitive measures are stipulated against any breach of the guaranteed confidentiality.

The important distinction between an anonymous and a confidential reporting system lies in the fact that, with an anonymous reporting system the reporter will submit unidentifiable reports. An anonymous report offers no possibility to derive further facts in the investigation process. However, with a confidential system the reporter will submit their name, and can thus be contacted during the investigation process for further clarification and feedback purposes.

The most important elements in establishing a new reporting culture are the following, which will be dealt with in the following headings:

- Legal framework
- Company commitment to safety
- Clear and unambiguous directions for reporting and accessibility to reporting means
- Professional handling of investigation and lesson dissemination
- Feed-back and knowledge sharing

However, it should be emphasised that this paper is not academic but practical in nature and that, while the reporting system and the process of implementing it described here are shaped by local circumstances, the author believes that the underlying practical and psychological mechanisms are universal for any safety critical business.

The legislative process in Denmark

In 2000, growing concerns about flight safety in Danish airspace were raised by the Danish Air Traffic Controllers Association. The concern was associated with losses of separation between aircraft that was not being reported due to the fear of sanctions of the reporter, particulary if he/she was partly or fully responsible for the incident. A fear that was real, since controllers previously had been prosecuted for such actions. Furthermore, the Danish press had during that period been dealing aggressively with apparent breaches of flight safety within certain airlines. These two factors- punishing Air Traffic Controllers with fines or license suspension and a biased focus by the press on aviation safety issues - had the effect of reducing the reporting of incidents. The whole aviation system in Denmark suffered from this, and no lessons were being learnt and disseminated from these events.

It should be added, however, that prior to 2000, the “culture of reporting” in Denmark was comparable to most northwest European countries – some occurrences did become reported, but there was an acknowledgement that “under-reporting” was being practiced. In contrast, in Denmark’s neighbouring country, Sweden - which has approximately the same amount of civilian air traffic - the number of flight safety occurrences reported was considerably larger than in Denmark.

Then, in 2000, in order to push for a change the Chairman of the Danish Air Traffic Controllers Association decided to be entirely open about the then current obstacles against reporting. During an interview on national television, she described frankly how the then current system was discouraging controllers from reporting. The journalist interviewing the ATCO chairman had picked up observations made by safety researchers that, as described above, Denmark had a much smaller number of occurrence reports than neighbouring Sweden. Responding to the interviewer’s query why this was so, the ATCO chairman proclaimed that separation losses between aircraft went unreported simply due to the fact that controllers - for good reasons - feared for retribution and disclosure. Moreover, she pointed out, flight safety was suffering as a consequence of this! These statements, broadcasted on a prime time news program, had the immediate effect that the Transportation Subcommittee of the Danish Parliament asked representatives from the Danish Air Traffic Controllers Association to explain their case to the Committee. Following this work, the Committee spent several of their 2000-01 sessions exploring various pieces of international legislation on reporting and investigation of aviation incidents and accidents. As a result of this, in 2001 the Danish government proposed a law that would make non-punitive, strictly confidential reporting possible [1].

The law would grant freedom from prosecution, even though the reporter had committed an erroneous act or omission that would normally be punishable. Furthermore the reports from this

scheme would be granted exemption from the provisions of the freedom of information act. Investigators would, by law, be obliged to keep information from the reports undisclosed. However the law would grant no immunity if gross negligence or substance abuse was present in the reported situations, and it would also be punishable by fine, **not** to report an incident in aviation.

In most democratic countries, the freedom of information act is an almost sacred institution. This fact is also the case in Denmark. It was acknowledged by the politicians and aviation specialists, that the public has a right to know the facts about the level of safety in Danish aviation. In order to accommodate this it was written in the law that the regulatory authority of Danish aviation, based on the incoming reports, should publish overview statistics two times per year, based on de-identified data from these reports.

This law was passed unanimously by the Danish parliament in May 2001[2]. Compared to other legal norms in Denmark, and probably in most countries, this law is unique. It is unique in the sense that it is the only law in Denmark that guarantees immunity from prosecution when an otherwise punishable offence has been committed.

During the legislative process, the public interest in the matter was surprisingly low and apart from a few editorials in national newspapers, the matter was not commented upon. After the regulatory authority, based on incoming flight safety reports, made their first statement, the public interest increased. However, the main interest in most media was not in the system itself, but in the aparent unsafe nature of Danish aviation!

The implementation process

After the law was passed, the Danish Aviation regulatory authority body, Statens Luftfartsvæsen, carried out the implementation of the regulatory framework. The regulatory authority subsequently issued instructions to the following groups:

- Pilots holding an Air Transportation Pilots License
- Air Traffic Controllers
- Certified Aircraft Mechanics
- Certified Airports.
- Pilots holding a General Aviation Pilots Licence.

For these five categories of license holders it would be mandatory to follow the reporting system.

Since both pilots and air traffic controllers have now to report various situations according to the reporting system, it is obvious that these two categories will sometimes be reporting situations basically created by the other. This will not incriminate either, as long as each professional abides by the obligation of reporting. This means that for example a situation created by air traffic control, reported by a pilot, will not incriminate the controller as long as the controller reports the same situation.

In order to make it clear which situations these personnel were obliged to report, the regulatory authority passed guidance material to each of the five categories. Since the situations that could pose a threat to aviation are different for the five categories, each of the five categories have their own set of descriptions of the mandatory reportable situations [3-4]. In the following sections, only the material and the process concerning Air Traffic Control will be dealt with.

Reporting and assessment of Safety Occurrences in Air Traffic Management

For Air Traffic Control the regulatory authority issued reporting categories that were derived from the EUROCONTROL standard ESARR 2. ESARR is a set of regulatory standards that has to be followed by aviation regulatory bodies throughout Europe [5-6].

ESARR 2 deals with Reporting and Assessment of Safety Occurrences in Air Traffic Management. To illustrate some of the reportable occurrences in Air Traffic Control, examples of these categories are mentioned in the following list:

- Separation losses between aircraft where no avoiding action was carried out
- Inadequate separation(where no minima exists)
- Runway incursions
- Aircraft deviation from clearances
- Deviation from procedures
- Failure in communication function
- Failure in surveillance function
- Failure in dataprocessing- and distribution function

Implementation in Denmark

Within Naviair (the Danish Air Traffic Control service provider employing all Air Traffic Controllers in Denmark), a high level decision was made to actively support the implementation process of this new reporting system. This decision was not made solely because it was mandatory, but because management foresaw a benefit for the company's main product *flight safety*. As a consequence of this, every Air Traffic Controller received a letter from management, explaining the new system stating Naviair´s commitment to enhance flight safety through the reporting and analysing of safety related events. The incident investigators, who was responsible for the implementation of the new system, were given the task of communicating the change, and were also given a full mandate and support by management.

An extensive briefing campaign was carried out in order to give information to every Air Traffic Controller about this new system. In the briefing process the controllers expressed many concerns, particularly pertaining to confidentiality and the non-punitive issues. These concerns were due to the existing culture and all anticipated. Questions were asked such as:

- Can we trust this new system?
- What will it be used for?
- Why more non-productive paperwork?
- We just handle the situations, so why report them?

These questions were typical and were asked by the controllers during the implementation process. They were dealt with by explaining the intentions of the law governing the reporting system; the law that would grant media and others no access to the reports, and the law that would secure freedom from prosecution. Furthermore it was emphasised that no major enhancement of flight safety would possible if no knowledge of the hazards was gathered and disseminated. It was explained to the controllers, that the reporting system could ultimately be the system that would be able to explain and hopefully eliminate the flaws that everybody recognised in everyday operation. We basically asked the Air Traffic Controllers to trust us, and take ownership of flight safety. In return we would try to deal effectively with flight safety.

The results

The reporting system started to operate on the 15th of August 2001. During the first 24 hours after starting, Naviair received 20 reports from Air Traffic Controllers! One year after the reporting system was started Naviair had received 980 reports-compared to the previous year´s 15 reports.

Still, the numbers from the new and the old 12-month period cannot be compared directly. With the new reporting system Air Traffic Controllers became obliged to report instances that were not compulsory to report beforehand. So the best comparison of the change would then be to compare the amount of reports for **losses of separation between aircraft** (they were mandatory reportable occurrences before implementation of this new system). The comparison is fair and informative and it serves to show the quite dramatic change in reporting culture, not least because these situations were the ones that Air Traffic Controllers were punished for beforehand.

Before the implementation of the reporting system *only* separation losses between aircraft were reported. These would average approximately 15 a year and two years after implementation 40-50 separation losses were reported per year.

It is important to mention that any company management that puts a system like this in place has to prepare for new and maybe unpopular knowledge. It may come as a surprise for the management of any company when more breaches of safety are being reported. It is very important that this new knowledge is not seen as a sign that safety is sliding. Rather it should be interpreted as an uncovering of things that have existed and gone unreported for years. The paradox remains, however, that the safest companies will initially be viewed as the unsafe companies due to their willingness to elicit a greater number of reports. For the time being it takes courage to be safe!

Investigation

The investigation process is one of the most important parts of a safety culture. It is of utmost importance that a company that puts a confidential non-punitive reporting system in place has to be professionally prepared to handle the challenge, and a formal process has to be set up to handle the reports.

The reports (they had to be submitted within maximum 72 hours) that were recieved in Naviair have varying content, ranging from small deviations or technical malfunctions, to serious losses of separation. Naturally, not all situations will recieve the same amount of attention and interest from the investigators.

In order to gain maximum flight safety benefit we have set up priorities for how we deal with the reports. In general, all reports are evaluated. The evaluation tries to establish whether immidiate correction is required. These situations would typically be cases of separation lossese between aircraft or serious procedural or technical issues.

All separation losses between aircraft will be investigated thoroughly. These incidents would be categorised and include the following:

- Separation minima infringement
- Runway incursion where avoiding action was necessary
- Inadequate separation between aircraft

The investigation will include gathering of all factual data such as voice recordings, radar recordings and the collection of flight progress strips, etc. After the factual data has been collected and analyzed the investigator will carry out interviews face to face with the involved controller(s) and other personnel relevant to the situation. The interview will be carried out with a human factors focus based on the HEIDI[a] taxonomy developed by EUROCONTROL. When the data gathering and interviews are completed the investigator will produce a written report on the incident, and the report has to be completed within maximum 10 weeks. The ultimate purpose of the report will be to recommend changes to prevent similar incidents.

In Naviair, the incident investigators have recieved training in both investigation techniques and human factorss and they are generally maintaining required to maintain their operational status, which has proven useful for keeping up credability with the controllers. Furthermore, it is recognized that it is not possible to produce a meaningfull report of an incident without current knowledge of air traffic control operations.

The form of the final report on incident follows the same format in every investigation. The report describes the fatual circumstances and contains the investigators" assessment of the following elements:

- Aircraft proximity and avoiding manouveres
- Safety nets - their impact on and relevance for the incident
- System aspects
- Human factors
- Procedures
- Conclusion
- Recommendations

In order to evaluate the effects of the reporting system it is interesting to look into the content of the incoming reports and the effect the investigation of these reports has had.

Example 1: Shortly after the reporting system was implemented, a tragic accident occurred at Milan airport in Italy. A Scandinavian Airlines MD 80 collided with a cargo terminal as a result of a collision with a Cessna (a small corporate aviation jet) on the runway. The collision happened because the Cessna had entered the Runway without clearance from the Tower, a so-called Runway Incursion. The preliminary investigation by the Italian Aviation Investigation Board demonstrated major flaws in Airport structure (signs and lighting), the handling of the situation by the Air Traffic Controller and Cessna pilot and the procedures in place at the time.

The accident naturally prompted a lot of attention in Scandinavia, since the MD 80 was an aircraft of the Scandinavian flag carrier and carried many Scandinavian passengers. After the accident, Naviair and the safety regulatory authority made an assessment of the new reporting system. The assessment was made in order to analyse, if any Runway incursions had been reported in Danish Air Traffic Control. It turned out that at the time of the accident, 40 Runway incursions had been reported through the system! These Runway incursions could be called "free of charge", since nothing happened as a result of them; but still there was a lot to learn. Immediately after this discovery, Naviair established a Runway Safety Task Force. The Task Force was asked to look into the nature of Runway incursions in Denmark. The Runway Safety Task Force was also asked to suggest recommendations to minimise the hazards etc. The Task Force work discovered that Danish aviation also had airports with ambiguous signs and lighting,

[a] HEIDI; Harmonisation of European Incident Definitions Initiative for ATM [7]

procedures that should be changed to minimise hazards etc. Each of these conditions were, as far as possible, corrected in accordance with the Task Force recommendations.

It is fair to assume that the work that was undertaken by The Runway Safety Task Force could not have been undertaken effectively without the reports from the controllers. These reports and the analysis of their content provided us with a number of conditions that deserved to be looked into; they even sometimes offered us causal factors to work with.

Example 2: Reports on incidents from air traffic controllers and pilots have highlighted human errors and the need for mitigating their consequences as one of the most important flight safety issues. In the Danish reporting system a relatively large number of reports about separation losses between aircraft have been received, and each occurrence has been subjected to a thorough investigation. What the investigations have shown is that human error (slips or lapses, misuse of procedures, bad procedures, interface between operator and machine etc.) account for 80-90 percent of the causal factors. In fact, this proportion does not reveal anything new, since the role of human error in accidents has been well known for years in aviation - as in all other safety critical industries.

What was new to us in the incident investigation unit was perhaps the insight that human error cannot be prevented. We needed to focus more on this fact, rather than trying to solve the human error puzzle. Therefore, we decide it was very important to focus on reducing the consequences of these errors [9]. Of course, considerable efforts have been made to eliminate all latent safety-threatening conditions, before a new procedure or system was put in place. But experience has showed that even the most rigorous safety assessment of a procedure or a system, cannot identify every latent condition, nor can it reveal every condition that will arise when you mix humans into the equation. Therefore it is of utmost importance for flight safety, that an effective feedback system is in place. The operator (Air Traffic Controller/Pilot) of the system or procedure can then use this system to report operator observed hazardous conditions. The analysis of these reports then serves to initiate corrections or the dessimination of information if needed. It is our experience that our new reporting system has proven its usefulness in this regard.

What the amount of, often self-incriminating, reports show is that a marked change in culture has taken place. Still, after two years it would also be unrealistic to think that all situations are reported. This is due to the fact that the reportable categories still need some time to be imbedded with the Air Traffic Controllers. Also the period before the time of implementation of the non-punitive scheme has engendered an atmosphere of distrust that takes time to overcome.

Flight Safety Partnership.

Another flight safety enhancing element that has offered itself after the new reporting system was implemented, is the sharing of flight safety knowledge. As a result of the investigations of the incoming reports, Naviair quickly realised that we in Air Traffic Control cannot handle flight safety alone. Many potential hazardous situations between aircraft arise as a consequence of the interface between Air Traffic Controllers and Pilots (misuse of phraseology, different understanding of procedures, different expectations etc). If we shall hope to make any new breakthrough in flight safety, it will be important to look at flight safety as a mutual process.

In order to deal more effectively with flight safety, Naviair decided to establish a Flight Safety Forum. Naviair subsequently invited flight safety officers from all the major Danish airlines to participate in discussion and knowledge sharing of flight safety relevant information. Everybody involved accepted this invitation and, as a result of this, we meet twice a year and address

operational flight safety in the Danish Airspace. Furthermore we have decided to share this information to be used in incident investigation.

Prerequisites for Reporting

It is recognised that a solid reporting culture relies mainly on the following:

- Trust/Confidentiality
- Non-punitive nature
- Ease of reporting
- Feedback to reporters
- Safety improvement

Trust/Confidentiality

It is of great importance that the reports are handled in a strictly confidential and trustworthy manner. It would be absolutely devastating to a reporting system if mass media had access to the reports. This can be illustrated by an example from Sweden.

In Sweden a reporting system had been in use for years in the Air Traffic Control system. The reporting system was used by Air Traffic Controllers to report any deviation from operational standards. The system was run on the basis of trust since the laws and the regulations underlying the system do in fact stipulate that Air Traffic Controllers may be prosecuted on the basis of the reports they submit. Furthermore, reports received from Air Traffic Controllers are not exempt from the freedom of information act and they may therefore be freely used and cited by the press. In the late 90`s the media and others had shown an increasing interest in the content of these reports (approxemately 1000 reports are recieved each year, only a small proportion of which deal with critical incidents). The media had asked the regulatory authority (Luftfartsverket) for information when Aviation Safety events had occurred. However, the regulatory authority had successfully convinced representatives from the media that the reporting system important to flight safety would suffer if the media were to take information from the reports and disclose it to the public.

Then, In 2000 an incident happened in Swedish airspace in which an aircraft was hit by lightning and requested a priority landing. The aircraft transmitted PAN PAN on the frequency indicating that a threatening situation was present. Due to a misunderstanding between the pilot and the Air Traffic Control unit handling the aircraft, the phraseology PAN PAN was misinterpreted. As a consequence of this the aircraft had to declare an emergency (MAYDAY), in order to be understood and get priority by Air Traffic Control.

News of this event was picked up by a national television network, which, appealing to the freedom of the press laws, obtained a copy of the voice recording of the communication between the aircraft and the control tower. The voice recording was then broadcast on television in a news programme describing a "failure" of Swedish Air Traffic Control leading to an alledgedly highly dangerous situation. In fact, the situation was not dramatic, but the replay of actual voice recording on television naturally caused uproar among the controllers in Sweden. Thankfully, the reporting system, which had been in place for more than 15 years, came out relatively well by the event. This was probably due to the fact that the system was embedded solidly within Swedish Air Traffic Control system. Still, this episode serves to illustrate that it takes only a few similar events to destroy confidence in an otherwise well-functioning reporting system.

What the above-mentioned examples highlight is that the reports have to be handled with care. In Naviair the reports are recieved by the watch supervisor on duty. He or she will place the report in a locked compartment to which only the Safety Investigators have access. Thus, the name of the person submitting the report will be known only to the Incident Investigators, and cannot be disclosed to others except under a very few and explicitly defined circumstances. The only conditions under which the Incident Investigators will reveal the name of a reporter to management are the following:

- Proficiency issues (i.e, when action is required due to evidence of diminished competence)
- Gross negligence (i.e., when described actions involve direct repudiation of duties)
- Substance abuse (i.e., alcohol or drugs)

Non-punitive nature

It is natural that Air Traffic Controllers and other aviation professionals, like everybody else in society, may not be expected to turn themselves in if they risk punishment; this reluctance to incriminate oneself is no doubt part of human nature. Therefore it is important for the quality of a flight safety reporting system that individuals, within certain well-defined limits, are granted immunity from sanctions. The immunity cannot, and shall not, be complete. It will always be necessary to punish individuals when they have been behaving in a grossly negligent way, and likewise substance abuse cannot be tolerated.

At the same time, experience from investigation show that gross negligence and substance abuse are extremely rare factors in aviation incidents and accidents.

In order for any reporting system to be useful, particularly where it is expected that individuals are expected to report their own mistakes, it is important that information obtained by self-reporting is not used to prosecute the reporter. This would also be inconsistent with international law.

The first court trial that has relevance for the new reporting system was held in Denmark in late 2002. A general aviation pilot was tried for flying in an unsafe way - he took off on a flight bringing too little fuel and had to land his aircraft in a cornfield. The trial had started based on the pilot's own report of the incident. But it was recognised during the trial process, that the incident report from the pilot could not be used as basis for the trial. The pilot was sentenced to pay a fine, but the prosecution had to build the case based on facts other than those submitted by the pilot in his report.

As described above, when a reporting system is non-punitive, this means that no criminal action and no disciplinary measures will be undertaken against the reporter on the basis of information contained in reports submitted. However, this does not mean that reports may always be submitted without consequences. Our experience has shown that action by the employer can sometimes be necessary in order to ensure safety (retraining, limitations in the amount of working positions, de-certification etc.). But the important point is that such consequences may never be initiated with a penalizing or disciplinary purpose – rather, their purpose is to either ensure that the reporter is brought back to a level of competence required for his duties or that he is relieved of his duties in a dignified way accepted by himself and his colleagues.

Ease of reporting

To prevent Air Traffic Controllers from feeling reporting is a burden, it is important that it can be done fairly easy. Naviair is currently developing at database that will make it possible for every Air Traffic Controller to report electronically, wherever they are, as long as they have access to a computer.

Feedback to reporters

Feedback is another vital element in a healthy reporting culture. Many reporting systems have become obsolete because the issue of feedback was neglected. If the reporters do not see any results from their efforts, they will, over time, consider the system as another "paper pushing" exercise. Upon adoption of the new reporting system a new incident investigation department was set up in Naviair. Today the size of the department (6 investigators and recording specialists) makes it possible to give feedback to the reporter, whenever, first, a report is recieved and, second, the analysis of the event is concluded.

Once a reporing system is started, it is very important that the organisation is ready to handle reports. Naviair started the reporting system with only two investigators (the "old" way of doing things). However, we very quickly realised that this was not enough in order to handle the high volume of reports and ensure feedback to all the reporters. Feed back is now offered twice a year in which all Air Traffic Controllers, in groups, will receive a safety briefing and discuss the safety events that have been reported and analyzed. These briefings are supported by replay of radar recordings whenever possible. Naviair also produces four issues of a company Safety Letter, where information from the reporting system is passed to all the Air Traffic Controllers.

Safety improvement

It is worth repeating that the overall goal of the whole exercise of establishing a flight safety reporting systems is to improve flight safety. In turn, the value of these systems has to be viewed with regard to their effect on flight safety. This can sometimes be a difficult task to perform, as a prevented accident will never appear in any statistics.

When we examine the improvements or changes we have made in our system (machine/procedure/human) since we implemented the reporting system, it is obvious to us that improvements have been made. Before the implementation of the reporting system, many of the flight safety relevant observations were reported, but they were reported to different departments in our company, thus eliminating the advantage of focused information gathering and dissemenation.

Conclusion

Today we feel confident that the system we put in place 2 years ago is solidly founded within our Air Traffic Control system. We base this assessment on what we hear when listening to the discussions among controllers and support staff that take place on and off record as well as on the amount and content of the reports we recieve. Thus, events that beforehand were only discussed among those present at the time of the event are now reported and the findings disseminated to the benefit of others. As Ralph Waldo Emerson puts it "Learn from the mistakes of others, you´ll never live long enough to make them all yourself".

Of course the system has suffered difficulties. Sometimes, Air Traffic Controllers do feel blamed when they learn of the conclusion of an investigation. Equally, in the minds of the individual involved, a non-punitive confidential culture may appear as a general amnesty for every mistake made; but that is not the case. Most of the investigated incidents have had human mistakes as their root cause. That fact can be hard to be face up to; and in such situations it is important to confront the individual in a way that inspires proactiveness both for the organisation and the individual so that both will learn.

What made all this possible? First of all it is important that the legal framework is in place to run a reporting system. Even the most well meaning management will have problems to install trust if legal action can still be undertaken against employees. In Europe the European Commission is in the final stage of delivering a Directive [8] that makes it mandatory for the EU member states to establish non-punitive confidential reporting systems in aviation. It is to be hoped that all or at least most European nations, in a few years from now, can participate effectively in sharing flight safety knowledge; thus maintaining and enhancing flight safety.

Secondly, the management of any company in a safety critical business, be that aviation, medical care, power or the nuclear industry etc. has to be committed. Safety starts at the top.

In order to give the Air Traffic Controllers themselves the ownership to flight safety, it is very important that the people that are communicating safety have a professional background. Many feelings become activated, and discussions will follow when you embark on the endeavour of communicating flight safety. These discussions and questions have to be answered by people who have "felt" the business themselves. Management will have to show support and be visible in the safety campaign, but the professional discussions have to be among professionals.

The ultimate test for any non-punitive confidential reporting system (the legal framework, the confidentiality, the psychology) will come if a country running such a system experiences an aviation disaster with loss of life. When this happens, everything takes a new and unknown course. To prepare for this it is important to focus on the fact that without aviation safety reporting systems, the likelihood of disasters are much greater.

References

1. Folketinget; Forslag til Lov om ændring af lov om luftfart (l128), (The Danish Parliament: Proposal for changing the law on aviation), 2001
2. Lov om luftfart (Danish Air Navigation order)
3. Civil Aviation Administration - Denmark; Regulations for Civil Aviation; BL 8-10; Regulations on mandatory reporting of flight safety occurrences; first edition; 2001
4. Civil Aviation Administration - Denmark; Regulations for Civil Aviation; BL 5-40; Order on the duty to report aircraft accidents and incidents; second edition; 1997
5. Eurocontrol; ESARR 2; Reporting and Assessment of Safety Occurrences in ATM; first edition; 1999
6. Eurocontrol; ESARR 2; GM1; Severity Classification Scheme for Safety Occurrences in ATM; first edition; 1999
7. Eurocontrol; SAF.ET1.ST02.1000-REP-10-00; HEIDI taxonomy, first edition; 2000
8. Commision of the European Communities (2000); Proposal for a Directive of the European Parliament and of the Council; COM(2000) 847 final; 2000/0343 (COD)
9. Thomas Bove (2002). Development and validation of human error management taxonomy in air traffic control. Risø-R-1378(EN) Risoe National Laboratory, 4000 Roskilde, Denmark (http://www.risoe.dk/rispubl/SYS/ris-r-1378.htm)

Madsen, M.D., A study of incident reporting in air traffic control - Moral dilemmas and the prospects of a reporting culture based on professional ethics. In: Investigation and reporting of incidents and accidents. IRIA 2002, Glasgow, 17-20 June 2002. GIST Technical Report, G2002-2) p. 161-170

Biography

Peter Majgård Nørbjerg, Head of incident investigation, Naviair, Naviair Alle 1, 2770 Kastrup, Denmark. (Email; pmn@naviair.dk; Phone; 0045 3247 8216)

The author is the Head of Incident Investigation with the Danish Air Traffic Control service provider Naviair.He has been an Air Traffic controller for 15 years in Tower and Approach positions at Scott Air Force Base Illinois, Sonderstrom Air Force Base Greenland and Copenhagen Airport Denmark. He became an Aviation Human factors specialist in 1996, and has served in a number of aviation safety related committees including the chairmanship of a Human Factors Subgroup for the Director General of Civil Aviation Regulation, and has been a participant to various human factors working groups within EUROCONTROL. He is a former Deputy Chairman of the Danish Air Traffic Controllers Association and in that capacity took part in the political process that led to the change of the law that governs flight safety reporting in Denmark. As an Incident Investigator he was responsible in 2001 for the implementation of a non-punitive confidential reporting system in the Danish Air Traffic Control system. [1]

[1] The author gratefully acknowledge a number of critical comments and suggestions from Anne Isaac (NATS) and Henning Boje Andersen (Risoe national laboratory) to the draft version of this paper

Should Reporting Programmes Talk to Each Other?

M. J. O'Leary; Humanautics; Peppard Common, Oxfordshire, UK

Keywords: Air Safety, Human Factors, Reporting Programmes

Abstract

British Airways employs two self-report programmes through which safety issues are communicated to the Safety department and Flight Operations. The primary channel, the Air Safety programme, collects data on technical, environmental, operational and crew issues. The ASR programme is an open reporting system managed by Flight Operations and its database holds the original reports, including crew names, and records whatever actions were undertaken. Naturally enough, whereas crew are more than happy to report technical or environmental problems, human nature makes crew more reticent in reporting issues when crew may have under-performed. The secondary reporting vehicle is the Human Factors Report programme that focuses on human performance and the factors that help or hinder it. The programme is voluntary and confidential and is managed by Safety Services. No crew names are recorded and no individual reports are published. The Safety department communicates relevant issues to Flight Operations. Here the two programmes and their interrelationship is described. A comparison is made of how each programme analyses the issues involved in the 'go-around' manoeuvre. It is argued that the power of the Air Safety analysis combined with Human Factors analysis is a more powerful tool than the simple sum of the two parts.

Introduction

Collecting data on safety failures and successes, collating and learning from this data, and applying this knowledge towards the improvement of safety are fundamental requirements in any industrial organisation. This is particularly true in high risk, high technology industries such as the nuclear, chemical and aviation industries. Ignoring these requirements results in the inability to manage safety effectively as those concerned will be unable to prioritise their risk management – even if they know which risks they face. Collecting the required information can be achieved in many ways. Internal reporting systems both automated and human are perhaps the most precise ways of safety data collection but it is also important to expand the focus and learn from other organisations, both within and without their own industries - and possibly from other departments in their own organisations. Learning from one's own mistakes is only bettered by learning from other peoples' mistakes!

Reason [1] eloquently and elegantly described the necessary feedback mechanisms required to establish effective safety feedback within an organisation (and, of course, such mechanisms can also be used for financial and commercial management) but what is sometimes overlooked is that Reason also recognised that multiple feedback loops were better than a single one. Amongst others, British Airways (BA) also realised this and over the last dozen or so years has developed a multiple loop feedback system for safety management. Some of these loops are self-report programmes but the system also includes automated flight parameter monitoring, safety process auditing, risk assessment, and maintenance monitoring and investigation programmes. Each individual feedback loop is embodied in a module of the British Airways Safety Information System, BASIS. Here I will focus on two self-report programmes used by flight crew. These are the Air Safety Reporting (ASR) and Human Factors Reporting (HFR) programmes. To draw the necessary comparisons between these two programmes I will use a 2002 study that examined how

well flight crew were managing ‘go-arounds’, a manoeuvre in which crew abort a landing at a late stage in the approach.

The Air Safety Reporting Programme

The ASR programme is the primary reporting vehicle for the passing of safety information from flight crew to flight operations management. The programme is mandatory and requires a report on any incident affecting air safety. It prescribes about 30 specific incident categories that must be reported and, moreover, it requires crew to report any incident that did or might influence air safety. Air Safety reports are written on a standard form which requests many specific details concerning the flight circumstances such as the time of day, the weight of the aircraft and precise details of the aircraft’s flightpath and position, as well as a (usually) short text description of the event. These data are stored in the ASR database. Analysts encode the reports with a small selection of BASIS References that characterise what kind of event had occurred, and also with a selection of BASIS Keywords that help describe the event more precisely. It should be noted that both References and Keywords are intrinsically negative, i.e., they indicate failures or factors that degrade safety. Below, this will be contrasted with the factors employed in the HFR analysis.

The References are largely high level causal categories such as ‘ATC’ or ‘Pilot Handling and Airmanship’. Keywords are used for lower level description of the events. Both References and Keywords can be used as keys to filter the database for specific types of events or issues. The frequencies of these can be graphically displayed over time or location or any one of a number of other factors. For instance, it might be required to examine the relative frequencies of go-arounds at a group of, or all, airports. This can be achieved with just a few keystrokes. Its ease of use allows accurate and rapid description and categorisation of all kinds of events and incidents. With approximately 8000 reports filed per year, experienced analysts can execute a risk assessment and make relevant and effective recommendations very speedily.

The ASR programme was the first of the many BASIS modules. Its success is largely due to its versatility. It includes basic filing cabinet functions such as storage and indexing; the facility to include analytic ‘keywords or ‘descriptors’ which also provides for a huge variety of search and filtering options; the search / filtering also supports a graphical system to indicate trends over time; and when networked (which is its normal mode) the built in communications processes provide an effective method of ‘actioning’ people and departments to investigate specific aspects of an event.

Another reason for the success of the programme, at least within BA itself, lies not in the technology but in the organisational culture in BA. The safety culture that supports such success results from hard organisational factors not (only) relying on the willing support of the flight crew. Successive CEOs have supported a vital corporate standing order that is directly concerned with the reporting of safety incidents. It states:

‘It is not normally the policy of British Airways to institute disciplinary proceedings in response to the reporting of any incident affecting safety.
‘British Airways will only consider initiating such disciplinary action where, in the Company's opinion, an employee has acted recklessly, or omitted to take action, in a way that is not in keeping with his/her responsibilities, training and/or experience.
‘The fact that the employee has fully complied with his/her responsibilities to report the circumstances and to co-operate fully throughout any investigation will weigh in his/her favour in the Company's consideration of the matter.

'However, in the event of an employee failing to report a safety related incident that they have discovered, they will be exposed to full disciplinary action.'

It is clear from the above that management considers that learning from incidents is more important than punishing the 'culprit', and that the real crime is not to report at all.

The Human Factors Reporting Programme

Both the HFR and the ASR programmes are worth papers to themselves but the latter is more complex than the former and therefore a more extended description of the HFR programme will follow. However, as it too will be relatively short the interested reader might learn more from O'Leary, Macrae & Pidgeon [2].

Whilst the ASR programme gives excellent information concerning what problems were affecting our flight crew there has generally been little feedback on WHY these problems occurred (particularly if the problem was caused by the crew!), or on how effectively the crew coped with them. Without the knowledge of problem cause and crew coping mechanisms, management's attempts at problem solving and anticipation were tentative. Consequently, a need was recognised for some form of proactive safety management tool and the human factors programme was introduced.

The HFR programme can be contrasted with the ASR in several ways. Unlike the ASR programme it is both confidential and voluntary due to the obvious sensitivity of reports that might frequently concern flight crew failures. Moreover it is managed by the Safety Services department independently of Flight Operations and run by line pilots who are specially trained in HF analysis. The issues raised in the reports are communicated to line management on a regular basis but great care is taken to separate the issues from the incidents in order to safeguard the identity of the reporters. The names of the reporters are not entered into the database.

When an ASR is filed, each crewmember of the originating flight receives a reply. If the ASR suggests that human factors might have been involved a human factors questionnaire accompanies the reply to the ASR. The HFR questionnaire elicits information with questions that mostly require descriptive answers. The questions are designed to help the reporter work through the incident quasi-chronologically and to help him or her recall the crew's actions and the reasons for their decisions and actions. The reply rate from solicited reports provides further useful information on about ten percent of the ASRs.

Human Factors Report analysis is complex in comparison with ASR analysis. The questionnaire focuses on Why the event occurred and How the crew solved or coped with the situation. Details from HFRs are entered only into the HFR database and can be supplemented with information from the related ASR and with information from a telephone or (occasionally) a face-to-face debrief with the reporter.

Each report is analysed with a set of 'Factors' concerning 'Crew Actions' and 'Influences' on those actions. The factors can be assigned in a negative - safety degrading - sense and, just as importantly, in a positive - safety enhancing – sense. Once these Factors are identified they are linked together to create an 'Event Sequence Diagram' (ESD) illustrating the flow of cause and effect throughout the incident. There are four groups of factors. The first reflects observable / describable crew behaviour or actions that can be defined as safe or unsafe. Three further categories apply to different influences on crew behaviour. The four are briefly described below.

Crew Actions are of three distinct types. The first concerns the activities of handling the aircraft and its systems, e.g., 'System Handling'. The second is based on the human error types described by Reason [1], e.g., 'Action Slip'. Third is the group of Crew Resource Management Teamskills (Helmreich, Butler, Taggart & Wilhelm, 1995). These describe a number of activities involved in the safe management of flight, e.g., 'Workload Management'.

Personal Influences describe the subjective feelings of physical and mental well-being, emotion, stress, motivation, and attention as described by the reporter. Examples are 'Boredom', 'Personal Stress', 'Tiredness' and 'Mode Awareness'.

Organisational Influences are those that are directly controlled by the company. Examples are 'Training', 'Technical Support', and 'Navigational Charts'.

Environmental Influences are those over which neither the reporter nor the company has any control. Examples are 'ATC Services', 'Technical Failure' and 'Weather'.

Crew actions differ from the influences in that they are generally observable and reportable. The majority of the influence factors are not so easily determined. In a few cases the influences can be inferred but it is essential that the inference is based only on evidence not assumption. This is particularly important in the assignment of the Personal Influences. These are subjective reports of personal feelings, states of arousal and attention. Assignment of any of the Personal Influences requires a direct report of these states by the reporter, not an inference by the analyst or by another crew member.

In the Factor assignment process Event Sequence Diagrams (ESD) are created for each report in a graphics image in the HFR database using a custom-built graphical interface. Analysts create the ESD by considering each action and influence and establishing all their interactions with the others. The final product normally represents a set of converging branches leading to an 'operational problem' of some sort and then terminating with one or more factors that indicate how the problem was solved (or not!). A very simplistic model of an ESD is shown below (figure 1). The arrows are the causal links between the factors. It is important to note that they are intended to indicate the direction of cause or influence, not just chronological relationships.

Figure 1 - Basic HF Event Sequence Diagram

The HFR and ASR programmes differ in several respects and O'Leary, Macrae & Pidgeon [2] gives a summary of many of the organisational differences (some not mentioned above). However, the major difference from a safety perspective is that the HFR programme was designed to elicit information about crew behaviour before, during and after an event, whereas the ASR programme was designed to elicit information concerning the event types and to quantify their relative frequencies. The next section will describe an investigation employing data from both programmes. Both sets of data are individually interesting and valuable but together they offer much more than just the sum of the parts. Relying only on one or other set would offer the safety analyst a much impoverished picture.

The go-around study

A 'go-around' is a manoeuvre in which the flight crew abort the landing at a late stage in the approach and for BA flight crew it is a requirement to file an ASR whenever a go-around is executed. In the first six months of 2002, BA crew reported 403 go-arounds through the Air Safety Reporting (ASR) programme. In the same period five years earlier, in 1997, the total was 440. Adjusting for the reduction of flight sectors across these years (approximately 6%) the go-around rate has reduced marginally by 2.5% over the five years.

Applying the Reference 'Go-around' as a filter to the ASR database will list the subset of reports in which a go-around is reported and, of course, the lists for 2002 and 1997 would include 403 and 440 reports respectively. Normally, analysis includes more than one Reference but it important to realise that the References do not necessarily have a causal relationship with the headline event, just that they were somehow associated with the event. A few more key presses can create a further list of all the associated References in frequency order. The two lists below in table 1 relate to the two different periods above and show the top ten BASIS References (excluding the Reference 'Go-around' itself).

Table 1 - BASIS References applied to Go-Arounds in the first six months of 1997 and 2002.

Jan – Jun 1997 G/As = 440		Jan – Jun 2002 G/As = 403	
WEATHER	152	AERODRM/LANDING SITE	125
AERODRM/LANDING SITE	132	WEATHER	114
ATC	81	ATC	77
PILOT HNDLG/AIRMNSHP	53	PILOT HNDLG/AIRMNSHP	73
GPWS	34	FLIGHT CONTROLS	22
FLIGHT CONTROLS	19	GPWS	20
AUTOFLIGHT	14	LANDING GEAR	11
LANDING GEAR	8	CABIN EQUIPMENT	4
CABIN EQUIPMENT	4	NAV EQUIPMENT	2
AIRPROX	1	FUEL	1

The table shows some interesting comparisons. Only two References, 'AUTOFLIGHT' and 'AIRPROX' disappear from the 1997 column and are replaced by 'NAV EQUIMPMENT' and 'FUEL' in the 2002 column. Excluding those References the others differ only slightly in ranking across a period of five years. It appears we still suffer and report the same problems as five year ago. Moreover, if we exclude 'WEATHER', which seems to have been considerably worse in the earlier period, two of the top three References, 'AERODROME/LANDING SITE' and 'ATC' appear to have very similar frequencies if the 6% reduction in sectors flown in the latter period is taken into account. However, contrary to this, 'PILOT HNDLG / AIRMANSHIP' has increased by a relative 50%. In 1997, this Reference accounted for 12% (53 of 440) of the total whereas, in 2002, this percentage had risen to 18% (73 of 403). This Reference is characterised in BASIS as 'Events where the handling or airmanship of the flight crew was a factor in the incident'. Thus the data above suggest only that flight crew might have been a causal factor in the go-around. This is neither a very surprising nor explicit conclusion. Knowing that the crew was a factor is not, by itself, very useful for implementing a training programme that might assist crew in avoiding go-arounds.

This short analysis shows the benefit of the ASR programme as we can pick on an issue, pull out the data and quickly execute a short analysis that can indicate whether the issue is deteriorating,

improving or just staying constant cross time – as in this case it appears to be. On the other hand we have not benefited much in terms of developing any useful idea of what are the real causes of go-arounds. Consequently we have gleaned little insight as to how we should go about implementing programmes directed towards reducing go-around frequency. A final aspect of the ASR analysis is that it is rare that note is taken of the occasions when the go-around itself was mismanaged. However, nearly three percent of crews reported that the go-around had been mismanaged in some way.

The present HF study took place against the background of renewed Flight Operations' interest in the go-around issue. When that work was undertaken the HF group in Safety Services undertook a short study to see whether we could extend and corroborate the ASR data described above. As the HFR programme offers a more sophisticated analysis of pilot performance than that available from the ASR programme, the HF analysis potentially offered a more detailed account of both the causes of go-arounds and of how well they were executed.

Human Factors Data Collection and Analysis

The return rate of the HF questionnaires is much lower than the ASR rate but the style of the programme and of the questionnaire elicits much franker and fuller disclosure of incident details than is normally obtained from an ASR. Consequently, in this study, the HFR programme's ability to elicit much more information on all aspects of an event potentially offered a much more thorough analysis than the above.

As previously described, both the actions and influences can be coded as safety positive or safety negative. In this study two lists of negative factors from each report were compiled. One list was for the factors that related to the flight immediately before the go-around was initiated and a separate list was composed of those relating to flight after the go-around was initiated. In this way the factors that had a causal role in the go-around could be analysed separately from those that resulted from the go-around. A similar analysis was applied to the safety positive factors.

The go-around HF reports were collected over the period between late April to early June 2002. A total of 132 HFR questionnaires were sent out covering 66 go-around incidents. The questionnaires were sent out with a covering letter explaining that this was a 'special' request for information for this go-around study. Fifty- four replies were received representing a return rate of just over 40%. This figure in itself is quite remarkable as it is over four times greater than the rate that would be expected from the normal operation of the HF programme. The 54 replies concerned 45 go-arounds. In nine cases reports were received individually from both the captain and the co-pilot involved in the same incident. As interest was primarily in the 'incident' rather than individual reports, when such 'paired' reports were received they were combined into a single incident analysis. Care was taken to eliminate double counting of factors when combining the reports concerning the same incident

From 45 incidents, 134 negative, pre go-around human factors were collected. The number of negative factors in each incident varied between one and ten as shown in Figure 2.

In the post go-around phase, shown in Figure 3, negative factors totalled 81 and the number in each incident varied between zero and eight indicating that 18 incidents had no post go-around internal or external disturbance. However, in 27 go-arounds some kind of problem had occurred It is interesting to note that 60% of the go-arounds did experience some internal or external disturbance contrasting dramatically with the less than three percent reported through the ASR programme.

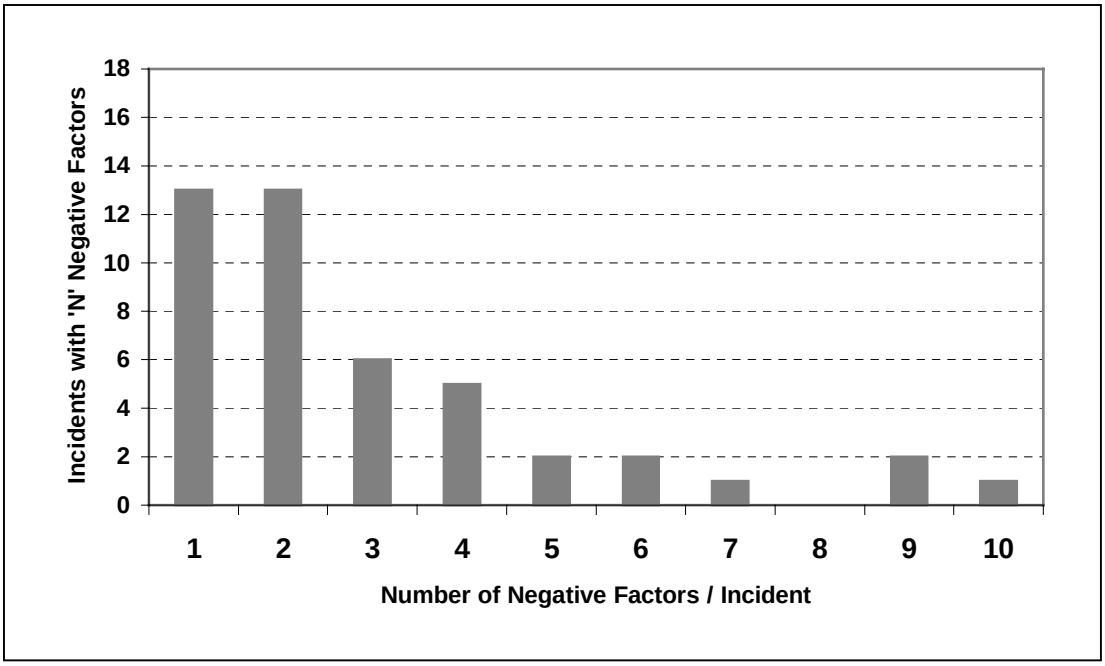


Figure 2 - Negative Factors per Incident Before Go-Around

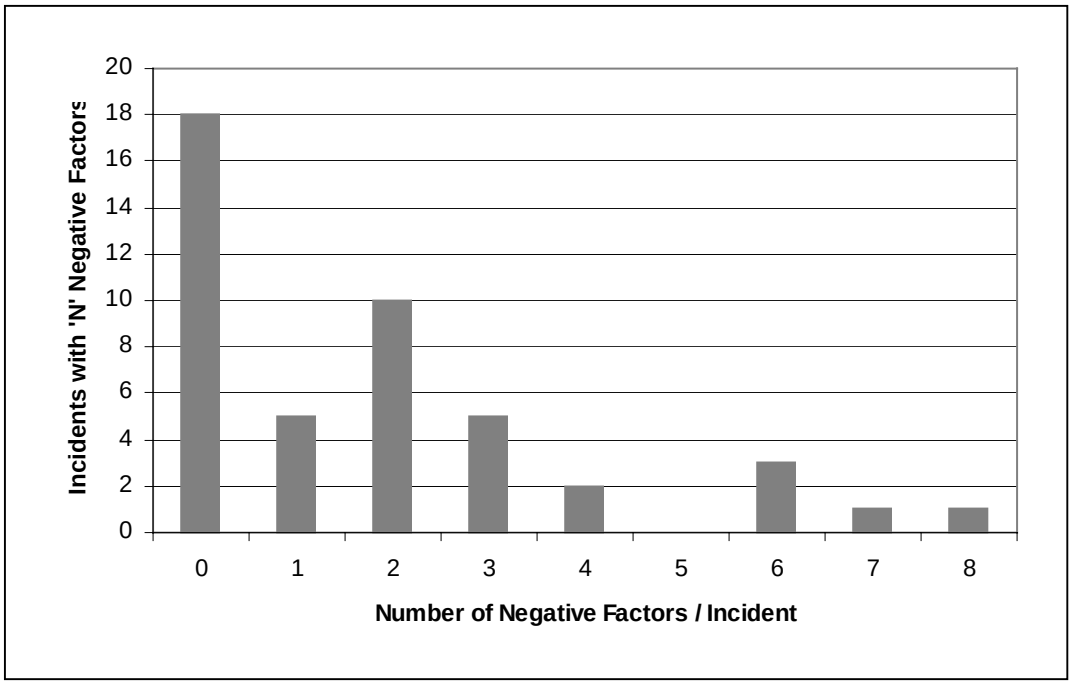


Figure 3 - Number of Negative Factors per Incident After Go-Around

Naturally, the most important aspect of this data is the identity of the negative factors in the analysis of the pre and post go-around phases. Table 2 below shows the 10 most frequently

assigned factors separately for both phases. 'N' is the number of assignments for any particular factor. The 'Total Factors' indicate the sum of all factors assigned, not just the ten most frequent factors shown here, and the 'Total Incidents' for the post-G/A phase differs from the pre-go-around phase as 18 go-arounds were untroubled.

Table 2 - Negative human factors applied to the pre and post go-around phases.

Rank	Pre Go-around	N	Post Go-around	N
1	ATC Services	28	Cross-Checking	11
2	Other Aircraft	22	Ops Stress	11
3	Met Conditions	13	ATC Service	8
4	Handling-Manual	8	Error	8
5	Airport Facilities	7	Handling-Manual	7
6	Prep / Planning	6	System Handling	5
7	Crew Comms	5	Prep/Plan	6
8	Mode Awareness	5	Currency	4
9	Ergonomics	4	Workload Management	3
10	Error	4	Training	3
Total Factors		134		81
Total Incidents		45		27

Comparing the Negative Factors before go-around, the left hand side of Table 2, with the BASIS ASR References in Table 1 shows a strong similarity between the top parts of the lists. 'WEATHER', 'AERODRM/LANDING SITE', 'ATC', 'PILOT HNDLG/AIRMNSHP' in Table 1 are directly comparable with 'Met Conditions', 'Airport Facilities', 'ATC Services' and 'Handling-Manual' in the pre go-around list of Table 2. The use of the factor 'Other Aircraft' in the same list indicates that another aircraft was somehow involved in the incident. None of the BASIS References or Keywords then represented the involvement of another aircraft although in the new ASR analysis 'Other Aircraft' has now been included. The similarity between the top part of the two lists is not surprising given that both describe causal factors in the go-arounds. (Perhaps one should use 'probable cause' in the ASR data).

The lower parts of the lists, however, differ markedly. Whilst Table 1 focuses on the technical causes of the go-arounds, the lower part of the pre go-around list in Table 2 represents mostly human failings. 'Prep / Planning', 'Crew Comms' and 'Mode Awareness' are the most common failings indicated by the HF analysis. The more general term 'Error'[1] is used to combine all the specific error types that can be recognised from the reporters' description of the event. This analysis offers a much clearer picture of the issues causing go-arounds than the single term 'Pilot Handling and Airmanship' in the ASR analysis.

Comparison between the pre and post go-around lists within Table 2 is even more interesting. The top two factors in the post go-around list, 'Cross-Checking' and 'Ops Stress', do not appear at all in the pre go-around list. Nor indeed do four other factors, 'System Handling', 'Currency', 'Workload Management' and 'Training'. In a general sense, the factors point to the effects of operational stress or overload, which frequently appears to be induced by ATC. This is aggravated by lack of practice, 'Currency' and 'Training', and poor 'Prep / Planning' and

[1] The term 'Error' used here is a simplification used to represent a variety of error forms. The error forms and their definitions are included in Appendix A along with the definitions of the other human factors used.

'Workload Management'. The consequence of these failings and pressures are under-performance in the handling of the aircraft flight path and configuration, and failures in 'Cross-Checking'. This latter factor heads the post go-around list and is not only the discipline of cross checking actions and communications with the other crew member but, more importantly in this case, is the requirement for the standard calls to be made in the approved manner at the correct time.

Until now we have focussed exclusively on the negative side of the analysis. There is, however, still another interesting story to tell and it involves the positive factors that are derived from the analysis. Table 3 below shows the human factors that either kept the flight safe or that recovered the situation after it had gone wrong. As before, 'N' is the number of assignments for any particular factor. The 'Total Factors' indicate the sum of all positive factors assigned and the 'Total Incidents' for the pre go-around phase differs from the previous 45 as no positive factors were assigned to this phase for seven incidents. For the post go-around phase only a few factors were assigned, 15 factors in eleven incidents. These factors were all crew actions that were directly involved in correcting problems that had occurred during the go-around.

Table 3 - Positive Human Factors applied to the Pre and Post Go-around Phases

Rank	Before Go-around	N	After Go-around	N	
1	Prep / Planning	25	Handling-Auto	4	
2	ATC Services	18	Handling-Manual	4	
3	Environment Awareness	15	Crew Comms	3	
4	Crew Comms	12	Assertiveness	2	
5	Mode Awareness	10	Role Conformity	1	
6	Handling-Manual	8	System Handling	1	
7	Currency	5			
8	Handling-Auto	5			
9	SOPs	5			
10	Workload Management	4			
Total Factors		118		15	
Total Incidents		38		11	

Positive factors before the go-around: Of the top ten positive factors applied to the before go-around phase, by far the most frequent was the crew action, 'Preparation / Planning'. This has often been promoted as the most important of the Teamskills and is the focus of the often quoted 'six Ps', i.e., Prior Preparation Prevents Poor Performance. 'Preparation / Planning' is examined in more detail below.

The second most frequent factor was 'ATC Services' indicating ATC's role in instructing go-arounds when spacing became less than necessary. Situation awareness factors were high on the list as was 'Crew Communications' both of which are, of course, fundamental to good flight management and safety.

Preparation and Planning: The teamskill, 'Preparation and Planning' accounted for more than 20% of all positive factors in the before go-around analysis. While studying the reports it was clear that good briefing and preparation before the event mostly led to a successful go-around. To objectify this possibility a further analysis was undertaken to establish whether a link between positive 'Preparation and Planning' and a positive go-around outcome could be established.

The negative factors after a go-around have already been presented in Table 2 and Preparation and Planning appeared in the top ten list. The list is mostly composed of various skill failures such as 'Cross-Checking' and 'Workload Management'. However, 'ATC Services', 'Training' and 'Currency' also appear in the list along with many other factors outside of the top ten such as 'Commercial Pressure', 'Ergonomics', 'Tiredness' and 'Airport Facilities'. Consequently, a link between positive 'Preparation and Planning' and, for example, the number of post go-around negative factors would not be a valid comparison. It was necessary to establish whether a link exists between 'Preparation and Planning' and the success or otherwise of the go-around itself.

The database was therefore sorted along two dimensions, 'Preparation and Planning' and go-around 'Outcome'. These two were divided into three categories, 'Positive', 'Negative' and 'Not Assessed'. The 'Not Assessed' was used when there was not enough information in the report to establish either Positive or Negative 'Preparation and Planning' or go-around 'Outcome'. Positive or negative 'Outcome' was determined on whether or not the go-around had been actioned without or with crew failure. The sorted data is presented in Table 4 below.

Table 4 - Matrix of 'Preparation & Planning' vs. Go-around 'Outcome' according to 'Positive', 'Negative' and 'Not Assessed' classification

		PREPARATION & PLANNING		
		Positive 27	Negative 11	Not Assessed 16
OUTCOME	Positive 32	23	1	8
	Negative 18	4	10	4
	Not Assessed 4	0	0	4

Interpreting Table 4 is not difficult. If 'Preparation & Planning' is positive then you have a likelihood of 23/27, 85%, of having a positive go-around outcome. Conversely if 'Preparation and Planning' is negative you have a ten to one chance of having a negative outcome. Other details in the table pale into insignificance in view of the above.

Summary

The go-around study combined the best of both worlds. The historic and statistical data from the ASR programme showed that the issue of go-around frequency and the major factors involved were unchanged over five years. With this starting point the application of human factors offered a fine-grain analysis of the issues and crew behaviour in the go-around scenario. Not only did it show where things were going wrong but also where some crew were being more effective than others in the application of teamwork and communications. This offers not only lessons to other crew but can also be used by training managers to implement effective training programmes.

I suggested in the introduction that organisational feedback loops are better in numbers rather than as singletons. Naturally the organisational structure and culture will define what precisely is

required but the very simple example offered here indicates that co-operative feedback loops can operate very effectively – together. Certainly in BA at least, there is a need for an ASR system that can amass huge numbers of reports and use them very effectively with statistical authority. It can show where real problems exist or where problems may be emerging by evaluating statistical trends in operational issues. There is also enormous value to be had in the human factors approach which, with a more precise and directed analysis process, can illuminate the important detail essential not only for understanding the problem but also for effectively specifying the effort required to reduce or eliminate the problem. Neither programme can do all these things entirely independently.

References

1. Reason, J. R. *Human Error*. Cambridge University Press, 1990.

2. O'Leary, M. J., Macrae, C., Pidgeon, N. F. *Safety Data Collection In British Airways Flight Operations*. Paper presented at IRIA 2002, Glasgow, 2002.

3. Helmreich, R. L., Butler, R. A., Taggart, W. R., Wilhelm, J. A. *Behavioral Markers in accidents and incidents: Reference List*. NASA/University of Texas/FAA Aerospace Crew Research Project. Technical Report 95-1, March, 1995.

Biography

M. J. O'Leary, Humanautics; Peppard Common, OXON, U.K.; telephone/fax - +44 1491628911; e-mail - mike@humanautics.com.

Mike O'Leary is a retired British Airways pilot and manager of the BA Human Factors reporting programme. His primary interests include the Human Factors of Aviation Safety, and sailing.

Appendix A: Definitions of Factors used in Table 2

FLIGHT CREW ACTIONS
CREW COMMUNICATION: Communication on the aircraft was not effective in informing everybody (including ATC) of relevant operational decisions, uncertainties, intentions, actions and aircraft/system states. Informing other crewmembers of stress and overload are also important aspects of this topic.
CROSS CHECKING: Indicates that standard calls and cross-checks were omitted, ineffective or deficient..
HANDLING – MANUAL: Manual flight handling degraded flight safety. Manual handling is to be understood as the direct manipulation of aircraft flight path and configuration. This can be effected either through the use of normal flight controls or through FCU / APFD or FMS, however it should result in an immediate change of flight parameters or configuration. (This factor is used when use of manual or automatic control cannot be ascertained. See following two factors.)
PREPARATION/PLANNING: Indicates that tactical (i.e., short term) pre-flight or in-flight planning and preparations were ineffective, omitted or inappropriately abbreviated.
SYSTEM HANDLING: Indicates faulty handling of aircraft systems, e.g., mechanical or electronic, or strategic handling of flight control systems through a FMS.
WORKLOAD MANAGEMENT: A failure of workload distribution, task priorities, distraction avoidance
ERRORS
ACTION SLIP: Indicates that a correct action was planned but an incorrect action was carried out unintentionally. E.g., selecting one switch in the belief that it was another, not because of ignorance of switch location but from absent-mindedness or distraction.
MEMORY LAPSE: A planned action was unintentionally omitted. We can assume that drills, checklists and procedures are 'planned'. Forgetting to complete, for instance, the Before Takeoff checks is a lapse.
MIS-RECOGNITION: Perceptual misinterpretation of visual or auditory data. E.g., mishearing ATC clearance, misreading instruments.
MISTAKE: An action was carried out as planned but the plan was faulty.
MISUNDERSTANDING: Conceptual misinterpretation of information. E.g., fault misdiagnosis, misunderstanding of manuals or clearances.
INFLUENCES ON FLIGHT CREW ACTIONS
AIRPORT FACILITIES: Airport facilities such as lighting, navigational aids or jetty docking facilities, were of poor quality or design causing operational difficulty.
ATC SERVICES: ATC instructions were unhelpful, led to unnecessary workload, conflicted with reasonable expectations or created an unsafe situation.
CURRENCY: Under-performance due lack of recent practice, or unfamiliarity with an airfield.
ERGONOMICS: Design of controls, displays or systems made them unfit for their intended purpose. This factor can be used in the case of 'degraded information' from displays and warnings etc.
MET CONDITIONS: Any meteorological condition that caused an operational difficulty.
MODE AWARENESS: Poor awareness of aircraft configuration, flight and powerplant parameters, flight control system modes, and the dynamic (rate of change / time to go e.t.c.) aspects of all of these. The parameters include such aspects as attitude, speed, altitude, heading, distance / time to go, and selected / armed / acquire / hold modes and the state of FMS data input and flight planning functions.
OPERATIONAL STRESS: Stress causing operational difficulty because of high operational workload or poor workload management. E.g., difficult procedures and drills, high workload departures / arrivals, or everything happening at once because of poor planning or organisation.
OTHER AIRCRAFT: Indicates that another aircraft caused an operational difficulty (e.g., runway occupation).
TRAINING: Indicates a training deficiency has been reported.

Applying STAMP in Accident Analysis[1]

Nancy Leveson, Mirna Daouk, Nicolas Dulac, and Karen Marais;
MIT, Cambridge, MA, U.S.A.

Keywords: Accident analysis, systems theory models, systems dynamics

Abstract

Accident models play a critical role in accident investigation and analysis. Most traditional models are based on an underlying chain of events. These models, however, have serious limitations when used for complex, socio-technical systems. Previously, Leveson proposed a new accident model (STAMP) based on system theory. In STAMP, the basic concept is not an event but a constraint. This paper shows how STAMP can be applied to accident analysis using three different views or models of the accident process and proposes a notation for describing this process.

Introduction

Most accident investigation and analysis rests on the use of event-chain models, i.e., the accident causation is described as a chain of failure events and human errors that led up to the actual loss event. Such models are limited in their ability to handle system accidents (arising from dysfunctional interactions among components and not just component failures), software-related accidents, complex human decision-making, and system adaptation or migration toward an accident over time [1,2].

In response to the limitation of event chain models, models based on systems theory have been proposed for use in accident analysis (see, for example Rasmussen [3]). STAMP (Systems-Theoretic Accident Modeling and Processes) is one such model that has been recently proposed [2]. Previously, only a description of the theoretical model underlying STAMP has been published. This paper shows how STAMP can be used in accident analysis and suggests notations that might be appropriate for representing and communicating the process leading to the accident.

The next section briefly describes STAMP. Then its application to a complex socio-technical accident is illustrated by applying it to the bacterial contamination of a water system in Walkerton Ontario in May 2000 where 2300 people became ill (in a town of 4800) and seven died [4].

Brief Description of STAMP

Accident models based on system theory consider accidents as arising from the interactions among system components and usually do not specify single causal variables or factors [5]. In STAMP, accidents are conceived as resulting not from component failures, but from inadequate control or enforcement of safety-related constraints on the design, development, and operation of the system. Safety is viewed as a *control problem*: accidents occur when component failures, external disturbances, and/or dysfunctional interactions among system components are not adequately handled. In the Space Shuttle *Challenger*, for example, the O-rings did not adequately control propellant gas release by sealing a tiny gap in the field joint. In the Mars Polar Lander loss, the software did not adequately control the descent speed of the spacecraft—it

[1] This research was partially supported by NASA grant NAG2-1843 and NSF ITR grant CCR-0085829.

misinterpreted noise from a Hall effect sensor as an indication the spacecraft had reached the surface of the planet.

Accidents such as these, involving engineering design errors, may in turn stem from inadequate control over the development process, i.e., risk is not adequately managed in the design, implementation, and manufacturing processes. Control is also imposed by the management functions in an organization—the *Challenger* accident involved inadequate controls in the launch-decision process, for example—and by the social and political system within which the organization exists. The role of all of these factors must be considered in accident analysis.

While events reflect the *effects* of dysfunctional interactions and inadequate enforcement of safety constraints, the inadequate control itself is only indirectly reflected by the events—the events are the *result* of the inadequate control. The control structure itself, therefore, must be examined to determine why the controls were inadequate to maintain the constraints on safety behavior and why the events occurred—for example, why the designers arrived at an unsafe design and why management decisions were made to launch despite warnings that it might not be safe to do so.

Systems are viewed, in this approach, as interrelated components that are kept in a state of dynamic equilibrium by feedback loops of information and control. A system is not treated as a static design, but as a dynamic process that is continually adapting to achieve its ends and to react to changes in itself and its environment. The original design must not only enforce appropriate constraints on behavior to ensure safe operation, but it must continue to operate safely as changes and adaptations occur over time. Accidents then are viewed as the result of flawed processes involving interactions among system components, including people, societal and organizational structures, engineering activities, and physical system components.

STAMP is constructed from three basic concepts: constraints, hierarchical levels of control, and process models. These concepts, in turn, give rise to a classification of control flaws that can lead to accidents. Each of these is described very briefly here; for more information see [2].

The basic concept in STAMP is not an event, but a constraint. In systems theory and control theory, systems are viewed as hierarchical structures where each level imposes constraints on the activity of the level below it—that is, constraints or lack of constraints at a higher level allow or control lower-level behavior [6]. Safety-related constraints specify those relationships among system variables that constitute the non-hazardous or safe system states—for example, the power must never be on when the access door to the high-voltage power source is open; pilots in a combat zone must always be able to identify potential targets as hostile or friendly; and the public health system must prevent the exposure of the public to contaminated water.

Instead of viewing accidents as the result of an initiating (root cause) event in a series of events leading to a loss, accidents are viewed as resulting from interaction among components that violate the system safety constraints. The control processes that enforce these constraints must limit system behavior to the safe changes and adaptations implied by the constraints. This definition of accidents fits both classic component failure accidents as well as system accidents.

Besides constraints and hierarchical levels of control, a third basic concept in STAMP is that of process models. Any controller—human or automated—must contain a model of the system being controlled. Accidents, particularly system accidents, frequently result from inconsistencies between the model of the process used by the controllers (both human and automated) and the actual process state; for example, the software does not know the plane is on the ground and raises the landing gear or the pilot thinks a friendly aircraft is hostile and shoots a missile at it.

When there are multiple controllers and decision makers, system accidents may also involve inadequate coordination of control actions and unexpected side effects of decisions or actions, again often the result of inconsistent process models. For example, two controllers may both think the other is making the required control action. Communication plays an important role here. Leplat suggests that accidents are most likely in *boundary* or *overlap* areas two or more controllers control the same process [5].

Starting from this definition of accidents in terms of inadequate control over system development and operations, control flaws can be classified and used during accident analysis or accident prevention activities to assist in identifying all the factors involved in the accident:

1. **Inadequate Enforcement of Constraints (Control Actions)**
   - 1.1 Unidentified hazards
   - 1.2 Inappropriate, ineffective, or missing control actions for identified hazards
     - 1.2.1 Design of control algorithm (process) does not enforce constraints
       - Flaw(s) in creation process
       - Process changes without appropriate change in control algorithm (asynchronous evolution)
       - Incorrect modification or adaptation
     - 1.2.2 Process models inconsistent, incomplete, or incorrect (lack of linkup)
       - Flaw(s) in creation process
       - Flaws(s) in updating process (asynchronous evolution)
       - Time lags and measurement inaccuracies not accounted for
     - 1.2.3 Inadequate coordination among controllers and decision makers (boundary and overlap areas)
2. **Inadequate Execution of Control Action**
   - 2.1 Communication flaw
   - 2.2 Inadequate actuator operation
   - 2.3 Time lag
3. **Inadequate or missing feedback**
   - 3.1 Not provided in system design
   - 3.2 Communication flaw
   - 3.3 Time lag
   - 3.4 Inadequate sensor operation (incorrect or no information provided)

In the rest of this paper, we show how STAMP can be applied to accident analysis using three different views or models of the accident process. A water contamination accident is used as an example.

The Water Contamination Accident

The accident occurred in May 2000 in the small town of Walkerton, Ontario, Canada [4]. Some contaminants, largely E. coli O157:H7 and Campylobacter jejuni entered the Walkerton water system through a well of the Walkerton municipal water system.

The Walkerton water system was operated by the Walkerton Public Utilities Commission (WPUC). Stan Koebel was the WPUC's general manager and his brother Frank its foreman. In May 2000, the water system was supplied by three groundwater sources: Wells 5, 6, and 7. The water pumped from each well was treated with chlorine before entering the distribution system.

The source of the contamination was manure that had been spread on a farm near Well 5. Unusually heavy rains from May 8 to May 12 carried the bacteria to the well. Between May 13

and May 15, Frank Koebel checked Well 5 but did not take measurements of chlorine residuals, although daily checks were supposed to be made. Low chlorine levels are a sign contaminants are overwhelming the disinfectant capacity of the chlorination process. Well 5 was turned off on May 15.

On the morning of May 15, Stan Koebel returned to work after having been away from Walkerton for more than a week. He turned on Well 7, but shortly after doing so, he learned a new chlorinator for Well 7 had not been installed and the well was therefore pumping unchlorinated water directly into the distribution system. He did not turn off the well, but instead allowed it to operate without chlorination until noon on Friday May 19, when the new chlorinator was installed.

On May 15, samples from the Walkerton water distribution system were sent to A&L Labs for testing according to the normal procedure. On May 17, A&L Labs advised Stan Koebel that samples from May 15 tested positive for E. coli and total coliforms. The next day (May 18) the first symptoms of widespread illness appeared in the community. Public inquiries about the water prompted assurances by Stan Koebel that the water was safe. By May 19 the scope of the outbreak had grown, and a pediatrician contacted the local health unit with a suspicion that she was seeing patients with symptoms of E. coli.

The Bruce-Grey-Owen Sound (BGOS) Health Unit (the government unit responsible for public health in the area) began an investigation. In two separate calls placed to Stan Koebel, the health officials were told that the water was ``okay." At that time, Stan Koebel did not disclose the lab results from May 15, but he did start to flush and super-chlorinate the system to try to destroy any contaminants in the water. The chlorine residuals began to recover. Apparently, Mr. Koebel did not disclose the lab results for a combination of two reasons: he did not want to reveal the unsafe practices he had engaged in from May 15-17 (i.e., running Well 7 without chlorination), and he did not understand the serious and potentially fatal consequences of the presence of E. coli in the water system. He continued to flush and super-chlorinate the water through the following weekend, successfully increasing the chlorine residuals. Ironically, it was not the operation of Well 7 without a chlorinator that caused the contamination; the contamination instead entered the system through Well 5 from May 12 until it was shut down on May 15.

On May 20, the first positive test for E. coli infection was reported and the BGOS Health Unit called Stan Koebel twice to determine whether the infection might be linked to the water system. Both times, Stan Koebel reported acceptable chlorine residuals and failed to disclose the adverse test results. The Health Unit assured the public that the water was safe based on the assurances of Mr. Koebel.

That same day, a WPUC employee placed an anonymous call to the Ministry of the Environment (MOE) Spills Action Center, which acts as an emergency call center, reporting the adverse test results from May 15. On contacting Mr. Koebel, the MOE was given an evasive answer and Mr. Koebel still did not reveal that contaminated samples had been found in the water distribution system. The Local Medical Officer was contacted by the health unit, and he took over the investigation. The health unit took their own water samples and delivered them to the Ministry of Health laboratory in London (Ontario) for microbiological testing.

When asked by the MOE for documentation, Stan Koebel finally produced the adverse test results from A&L Laboratory and the daily operating sheets for Wells 5 and 6, but said he could not produce the sheet for Well 7 until the next day. Later, he instructed his brother Frank to revise the Well 7 sheet with the intention of concealing the fact that Well 7 had operated without a

chlorinator. On Tuesday May 23, Stan Koebel provided the altered daily operating sheet to the MOE. That same day, the health unit learned that two of the water samples it had collected on May 21 had tested positive for E. coli.

Without waiting for its own samples to be returned, the BGOS health unit on May 21 had issued a boil water advisory on local radio. About half of Walkerton's residents became aware of the advisory on May 21, with some members of the public still drinking the Walkerton town water as late as May 23. The first person died on May 22, a second on May 23, and two more on May 24. During this time, many children became seriously ill and some victims will probably experience lasting damage to their kidneys as well as other long-term health effects. In all, seven people died and more than 2300 became ill.

Looking only at these proximate events, it appears that this is a simple case of incompetence, negligence, and dishonesty by WPUC employees. In fact, the government argued at the Inquiry that Stan Koebel and the Walkerton PUC were solely responsible for the outbreak and that they were the only ones who could have prevented it. In May 2003, exactly three years after the accident, Stan and Frank Koebel were arrested for their connection to the loss.

A STAMP analysis, however, provides a much more informative and useful understanding of the accident and what might be changed to prevent future repetitions (besides simply firing or arresting the Koebel brothers). In fact, the stage for the accident had been set over a large number of years by actions at all levels of the socio-technical system structure—an example of how complex socio-technical systems can migrate toward an accident. In this case as in many others, degradation in the water safety control structure had occurred over time, without any particular single decision to do so but simply as a series of decisions that moved the public water system slowly toward a state of high risk where any slight error or deviation from the normal could lead to a major accident. Degradation of the safety control structure may be related to *asynchronous evolution* [5], where one part of a system changes without the related necessary changes in other parts. Changes to subsystems may be carefully designed, but consideration of their effects on other parts of the system, including the control aspects, may be neglected or inadequate. Asynchronous evolution may also occur when one part of a properly designed system deteriorates.

Vicente and Christoffersen [7] have used the Walkerton accident to test the explanatory adequacy of Rasmussen's framework for risk management in a dynamic society [3]. While the Rasmussen approach does add analysis at multiple organizational levels, it does this in essence by adding event chains at each level (physical, system, operator, management, government) with links between the chains. In this paper, we use the same accident to illustrate how a pure systems theory model, i.e., STAMP, can be used to analyze the Walkerton accident and to show how three views or models of the accident can be used to explain it.

The first step in creating a STAMP analysis is to identify the system hazards, the system safety constraints, and the control structure in place to enforce the system safety constraints, as shown in the next section. Each component of the socio-technical control structure will have safety constraints relevant to the particular functions of the component. Together, the safety constraints on all the components must be adequate to enforce the overall system safety constraints.

We show the dynamic aspects of accidents in two ways. The first shows the changes in the static safety control structure over time. These models are essentially a series of static snapshots of the control structure, and they do not show the dynamic processes in effect that led to the changes.

For the latter, we use system dynamics models. At this point in the analysis, it is possible to examine the proximate events and their relationship with the safety control structure.

The third modeling effort provides an overall explanation of the accident. This model contains a summary of the other models: for each of the control components, it shows the inadequate control actions and decisions and the factors (using the STAMP classification shown above) that led to the accident. This final summary model provides the information necessary to make recommendations to prevent future accidents arising from the same inadequate controls over safety.

The Socio-Technical Water Safety Control Structure

Figure 1 shows the basic Ontario water quality safety control structure. For space reasons, we consider only the changes to the safety control structure over time, but a complete root cause analysis of the accident would also need to consider the decisions made during the water system design that contributed to the accident.

The general system hazard related to the accident is public exposure to E. coli or other health-related contaminants through drinking water. This hazard leads to the following system safety constraint:

The safety control structure must prevent exposure of the public to contaminated water.

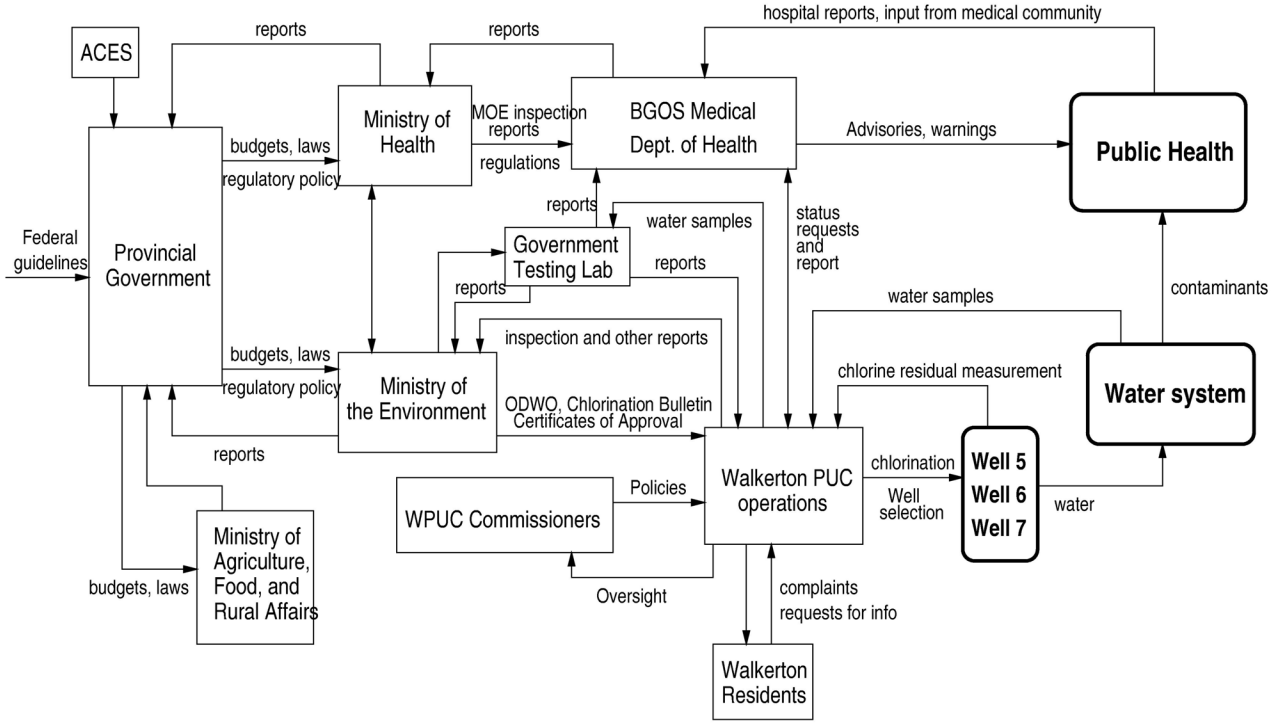
- *Water quality must not be compromised*.
- *Public health measures must reduce risk of exposure if water quality is compromised (e.g., boil water advisories).*

These general constraints must be enforced by requirements and constraints on the entire control structure. The Canadian federal government (not shown in the figure) is responsible for establishing a nationwide public health system and ensuring it is operating effectively. Federal guidelines are provided to the provinces, but responsibility for water quality is primarily delegated to the individual provinces.

The Ontario government is responsible for regulating and overseeing the safety of Ontario's drinking water. It does this by providing budgets for the ministries involved, in this case the Ministry of the Environment (MOE), the Ministry of Health (MOH) and the Ministry of Agriculture, Food, and Rural Affairs, and by passing laws and adopting government policies affecting water safety.

The Ministry of Agriculture, Food, and Rural Affairs is responsible for regulating agricultural activities with potential impact on drinking water sources. In fact, there was no watershed protection plan to protect the water system from agricultural runoff. Instead, the Ministry of the Environment was responsible for ensuring that the water systems could not be affected by such runoff.

The Ministry of the Environment (MOE) has primary responsibility for regulating and for enforcing legislation, regulations, and policies that apply to the construction and operation of municipal water systems. Guidelines and objectives are set by the MOE, based on Federal guidelines. They are enforceable through Certificates of Approval issued to public water utilities operators, under the Ontario Water Resources Act. The MOE also had legislative responsibility for building and maintaining water treatment plants.



Safety Requirements and Constraints :

Federal Government

- Establish a nationwide public health system and ensure it is operating effectively.

Provincial Government

- Establish regulatory bodies and codes of responsibilities, authority, and accountability
- Provide adequate resources to regulatory bodies to carry out their responsibilities.
- Provide oversight and feedback loops to ensure that provincial regulatory bodies are doing their job adequately.
- Ensure adequate risk assessment is conducted and effective risk management plans are in place.

Ministry of the Environment

- Ensure that those in charge of water supplies are competent to carry out their responsibilities.
- Perform inspections and surveillance. Enforce compliance if problems found.
- Perform hazard analyses to identify vulnerabilities and monitor them.
- Perform continual risk evaluation for existing facilities and establish new controls if necessary.
- Establish criteria for determining whether a well is at risk.
- Establish feedback channels for adverse test results. Provide multiple paths.
- Enforce legislation, regulations and policies applying to construction and operation of municipal water systems.
- Establish certification and training requirements for water system operators.

ACES

- Provide stakeholder and public review and input on ministry standards

Ministry of Health

- Ensure adequate procedures exist for notification and risk abatement if water quality is compromised.

Government Water Testing Labs

- Provide timely reports on testing results to MOE, PUC, and and Medical Dept. of Health

WPUC Commissioners

- Oversee operations to ensure water quality is not compromised.

WPUC Operations Management

- Monitor operations to ensure that sample taking and reporting is accurate and adequate chlorination is being performed.

WPUC Operations

- Measure chlorine residuals.
- Apply adequate doses of cholorine to kill bacteria.

BGOS Medical Department of Health

- Provide oversight of drinking water quality.
- Follow up on adverse drinking water quality reports.
- Issue boil water advisories when necessary.

Figure 1 - The Basic Water Safety Control Structure. Lines going into the left of a box are control lines. Lines from or to the top or bottom of a box represent information, feedback, or a physical flow. Rectangles with sharp corners are controllers while rectangles with rounded corners represent plants.

The Ministry of the Environment had two guidelines related to water safety. Note that guidelines, unlike regulations, are not legally binding. The Chlorination Bulletin required water systems to treat well water with sufficient chlorine to inactivate any contaminants in the raw water and to

sustain a chlorine residual of 0.5 mg/L of water after 15 minutes of contact time. The Ontario Drinking Water Objectives (ODWO) provided further guidelines on the operation of public water systems, including a requirement for the water testing laboratories, which were almost all government run, to report adverse test results directly to the MOE and to the local Medical Officer of Health (part of the MOH). The Medical Officer of Health could then decide whether to issue a boil water advisory.

The MOE was also responsible for public water system inspections and drinking water surveillance, for setting standards for certification of municipal water systems, and for continuing education requirements for operators to maintain competence as knowledge about water safety increased.

The Ministry of Health supervises local Health Units, in this case, the Bruce-Grey-Owen-Sound (BGOS) Department of Health, run by local Officers of Health in executing their role in protecting public health. The BGOS Medical Dept. of Health receives inputs from various sources, including hospitals, the local medical community, the Ministry of Health, and the WPUC, and in turn is responsible for issuing advisories and alerts if required to protect public health. Upon receiving adverse water quality reports from the government testing labs or the MOE, the local public health inspector in Walkerton would normally contact the WPUC to ensure that follow-up samples were taken and chlorine residuals maintained.

The public water system in Walkerton is run by the Walkerton Public Utilities Commission (WPUC), which operates the wells and is responsible for chlorination and for measurement of chlorine residuals. Oversight of the WPUC is provided by elected WPUC Commissioners. The Commissioners were responsible for establishing and controlling the policies under which the PUC operated, while the general manager (Stan Koebel) and staff were responsible for administering these policies in operating the water facility.

This then is the basic water safety control structure. The next step in the STAMP analysis is to examine the changes in this structure leading to the accident.

Changes in the Safety Control Structure Leading Up to the Accident

The water safety control structure started out with some weaknesses that were mitigated by the presence of other controls. As the other controls weakened or disappeared over time, the entire socio-technical system moved to a state where a small change in the operation of the system or in the environment (in this case, unusually heavy rain) could lead to a tragedy. Almost all the information about the accident that follows is from the official Walkerton Inquiry report [4] or from a magazine article about the tragedy by a local farmer [8]. Where possible the facts in each of these reports were checked with other sources.

Walkerton Well 5 was built in 1978 and issued a Certificate of Approval by the MOE in 1979. Despite potential problems—the groundwater supplying the well was recognized as being vulnerable to surface contamination—no explicit operating conditions were imposed at the time (*missing control action*). Well 5 was a very shallow well: all of its water was drawn from an area between 5m and 8m below the surface. More significantly, the water was drawn from an area of bedrock, and the shallowness of the soil overburden above the bedrock along with the fractured and porous nature of the bedrock itself made it possible for surface bacteria to make its way to Well 5.

Although the original Certificate of Approval for Well 5 did not include any special operating conditions, over time MOE practices changed (*asynchronous evolution*). By 1992, the MOE had developed a set of model operating conditions for water treatment and monitoring that were routinely attached to new Certificates of Approval for municipal water systems. There was no effort, however, to determine whether such conditions should be attached to existing certificates, such as the one for Well 5 (*missing control action*).

The ODWO was amended in 1994 to require the continuous monitoring of chlorine residuals and turbidity for wells supplied by a groundwater source that was under the direct influence of surface water (as was Walkerton's Well 5). Automatic monitoring and shutoff valves would have mitigated the operational problems at Walkerton and prevented the deaths and illness associated with the E. coli contamination in May 2000 if the requirement had been enforced in existing wells. However, at the time, there was no program or policy to review existing wells to determine whether they met the requirements for continuous monitoring (*control action omission*). In addition, MOE inspectors were not directed to notify well operators (like the Koebel brothers) of the new requirement nor to assess during inspections if a well required continuous monitoring (*missing control action*). Stan and Frank Koebel lacked the training and expertise to identify the vulnerability of Well 5 themselves and to understand the resulting need for continuous chlorine residual and turbidity monitors.

Operating conditions should theoretically have been imposed by the municipality, the Walkerton Public Utilities Commissioners, and the manager of the WPUC. The municipality left the operation of the water system to the WPUC (*inadequate control actions*). The WPUC Commissioners, who were elected, became over the years more focused on the finances of the PUC than the operations (*asynchronous evolution*). They had little or no training or knowledge of water system operations or even water quality itself (*inadequate mental models*). Without such knowledge and with their focus on financial issues, they gave all responsibility for operations to the manager of the WPUC (Stan Koebel) and provided no other operational oversight.

The operators of the Walkerton water system did not intentionally put the public at risk. Stan Koebel and the other WPUC employees believed the untreated water was safe and often drank it themselves at the well sites (*inadequate mental models*). Local residents also pressed the WPUC to decrease the amount of chlorine used because they objected to the taste of chlorinated water (*hazardous inputs, inadequate control*).

Although Mr. Koebel knew how to operate the water system mechanically, he lacked knowledge about the health risks associated with a failure to properly operate the system and of the importance of following the MOE requirements for treatment and monitoring. This incorrect mental model was reinforced when over the years he received mixed messages from the MOE about the importance of several of its own requirements.

Before 1993, there were no mandatory certification requirements for water system operators or managers. Stan and Frank Koebel were not qualified to hold their positions within the WPUC, but they were certified in 1993 through a grandfathering scheme based solely on experience. They were not required to take a training course or to pass any examinations (*missing and inadequate control actions*).

After the introduction of mandatory certification in 1993, the MOE required 40 hours of training a year for each certified operator. Stan and Frank Koebel did not take the required amount of training, and the training they did take did not adequately address drinking water safety. The

MOE did not focus the training on drinking water safety and did not enforce the training requirements (*missing control action*).

The Koebel brothers and the Walkerton commissioners were not the only ones with inadequate training and knowledge of drinking water safety. Evidence at the Inquiry showed that several environmental officers in the MOE's local office were unaware that E. coli was potentially lethal and their mental models were also incorrect with respect to other matters essential to water safety.

Without regulations or oversight or enforcement of safe operating conditions, and with inadequate mental models of the safety requirements, operating practices have a tendency to change over time in order to optimize a variety of goals that conflict with safety. In the case of Walkerton, this change began almost immediately. The Inquiry report says that many improper operating practices had been going on for years before Stan Koebel became manager. He simply left them in place. These practices, some of which went back 20 years, included misstating the locations at which samples for microbiological testing were taken, operating wells without chlorination, making false entries in daily operating sheets, failing to measure chlorine residuals daily, failing to adequately chlorinate the water, and submitting false annual reports to the MOE (*inadequate "actuator" operation, incorrect feedback*).

All of these weaknesses in the control over the Walkerton (and other municipalities) water quality might have been mitigated if the source of contamination of the water had been controlled. A weakness in the basic water control structure was the lack of a government watershed and land use policy for agricultural activities that can impact drinking water sources. In fact, at a meeting of the Walkerton town council in November 1978 (when Well 5 was constructed), MOE representatives suggested land use controls for the area around Well 5, but the municipality did not have the legal means to enforce such land use regulations because the government of Ontario had not provided the legal basis for such controls.

Walkerton is at the heart of Ontario's Bruce County, a major farming area. Whereas the existing water quality infrastructure and physical well designs were able to handle the amount of manure produced when farms typically produced 50 or 60 animals at a time, the increase in factory farms (each of which might have 1200 hogs) led to runoff of agricultural contaminants and put pressure on the drinking water quality infrastructure (*asynchronous evolution*). At the time of the accident, the county had a population of only 60,000 people, but had 163,000 beef cattle and 100,000 hogs. A single 1200 hog factory farm can produce as much waste as 60,000 people and the entire animal population in the county at that time produced as much waste as 1.6 million people. This animal waste is spread on the fields adjacent to the farms, which cannot absorb such massive quantities of manure. Contamination of the groundwater and surrounding waterways is the result. At the same time, the spreading of manure had been granted a long-standing exemption from EPA requirements (*inadequate control actions*).

Annual reports of the Environment Commissioner of Ontario for the four years before the Walkerton accident included recommendations that the government create a groundwater strategy. A Health Canada study stated that the cattle counties of Southwestern Ontario, where Walkerton is located, are high-risk areas for E. coli infections. The report pointed out the direct link between cattle density and E. coli infection, and showed that 32 percent of the wells in rural Ontario showed fecal contamination. Dr. Murray McQuigge, the Medical Officer of Health for the BGOS Health Unit (and the man who handled the Walkerton E. coli outbreak) warned in a memo to local authorities that ``poor nutrient management on farms is leading to a degradation of the quality of ground water, streams, and lakes." Nothing was done in response to these warnings (*ignored feedback*).

The control structure quickly started to degrade even further in effectiveness with the election of a conservative provincial government in 1995. A bias against environmental regulation and red tape led to the elimination of many of the government controls over drinking water quality. A Red Tape Commission was established by the provincial government to minimize reporting and other requirements on government and private industry. At the same time, the government disbanded groups like the Advisory Committee on Environmental Standards (ACES), which reviewed ministry standards including those related to water quality. At the time of the Walkerton contamination, there was no opportunity for stakeholder or public review of the Ontario clean water controls (*feedback loops eliminated*).

Budget and staff reductions by the conservative government took a major toll on environmental programs and agencies (although budget reductions had started before the election of the new provincial government). The MOE budget was reduced by 42% and 900 of the 2400 staff responsible for monitoring, testing, inspection, and enforcement of environmental regulations were laid off. The official Walkerton Inquiry report concludes that the reductions were not based on an assessment of the requirements to carry out the MOE's statutory requirements nor on any risk assessment of the potential impact on the environment or, in particular on water quality. After the reductions, the Provincial Ombudsman issued a report saying that cutbacks had been so damaging that the government was no longer capable of providing the services that it was mandated to provide. The report was ignored.

In 1996, the Water Sewage Services Improvement Act was passed, which shut down the government water testing laboratories, downloaded control of provincially owned water and sewage plants to the municipalities, eliminated funding for municipal water utilities, and ended the provincial Drinking Water Surveillance Program, under which the MOE had monitored drinking water across the province (*controls and feedback loops eliminated*).

The ODWO directed testing labs to report any indications of unsafe water quality to the MOE and to the local Medical Officer Of Health. The latter would then decide whether to issue a boil water advisory. When government labs conducted all of the routine drinking water tests for municipal water systems throughout the province, it was acceptable to keep the notification protocol in the form of a guideline under the ODWO rather than a legally enforceable law or regulation. However, the privatization of water testing and the exit of government labs from this duty in 1996 made the use of guidelines ineffective in ensuring necessary reporting would occur. At the time, private environmental labs were not regulated by the government. No criteria were established to govern the quality of testing or the qualifications or experience of private lab personnel, and no provisions were made for licensing, inspection, or auditing of private labs by the government (*inadequate controls*). In addition, the government did not implement any program to monitor the effect of privatization on the notification procedures followed whenever adverse test results were found (*inadequate control algorithm and missing feedback loop*).

At the time of privatization in 1996, the MOE sent a guidance document to those municipalities that requested it. The document strongly recommended that a municipality include in any contract with a private lab a clause specifying that the laboratory directly notify the MOE and the local Medical Officer of Health about adverse test results. There is no evidence that the Walkerton PUC either requested or received this document (*communication flaw*).

After laboratory testing services for municipalities were assumed by the private sector in 1996, the MOH Health Unit for the Walkerton area sought assurances from the MOE's local office that the Health Unit would continue to be notified of all adverse water quality results relating to

community water systems. It received that assurance, both in correspondence and at a meeting of representatives from the two agencies.

In 1997, the Minister of Health took the unusual step of writing to the Minister of the Environment requesting that legislation be amended to ensure that the proper authorities would be notified of adverse water test results. The Minister of the Environment declined to propose legislation, indicating that the ODWO dealt with the issue. On several occasions, officials in the MOH and the MOE expressed concerns about failures to report adverse test results to local Medical Officers of Health in accordance with the ODWO protocol. But the anti-regulatory culture and the existence of the Red Tape Commission discouraged any proposals to make notification legally binding on the operators or municipal water systems and private labs.

The testing laboratory used by Walkerton in May 2000, A&L Canada Laboratories East, was unaware of the notification guideline in the ODWO (*communication flaw*). In fact, they considered test results to be confidential and thus improper to send to anyone but the client, in this case, the WPUC manager Stan Koebel *(incorrect process model*). The MOE had no mechanism for informing private laboratories of the existing guidelines for reporting adverse results to the MOE (*missing control channel*).

Another important impact of the 1996 law was a reduction in the MOE water system inspection program (*degradation of feedback loop*). The cutbacks at the MOE negatively impacted the number of inspections, although the inspection program had other deficiencies as well.

The MOE inspected the Walkerton water system in 1991, 1995, and 1998. At the time of the inspections, problems existed relating to water safety. Inspectors identified some of them, but unfortunately two of the most significant problems—the vulnerability of Well 5 to surface contamination and the improper chlorination and monitoring practices of the PUC—were not detected (*inadequate actuator operation*). Information about the vulnerability of Well 5 was available in MOE files, but inspectors were not directed to look at relevant information about the security of water sources and the archived information was not easy to find (*inadequate control algorithm*). Information about the second problem, improper chlorination and monitoring practices of the WPUC, was there to be seen in the operating records maintained by the WPUC. The Inquiry report concludes that a proper examination of the daily operating sheets would have disclosed the problem. However, the inspectors were not instructed to carry out a thorough review of operating records (*inadequate control*).

The 1998 inspection report did show there had been problems with the water supply for years: detection of E. coli in treated water with increasing frequency, chlorine residuals in treated water at less than the required 0.5 mg/L, non-compliance with minimum bacteriological sampling requirements, and not maintaining proper training records.

The MOE outlined improvements that should be made, but desperately short of inspection staff and faced with small water systems across the province that were not meeting standards, it never scheduled a follow-up inspection to see if the improvements were in fact being carried out (*inadequate control, missing feedback loop*). The Inquiry report suggests that the use of guidelines rather than regulations had an impact here. The report states that had the Walkerton PUC been found to be in non-compliance with a legally enforceable regulation, as opposed to a guideline, it is more likely that the MOE would have taken stronger measures to ensure compliance—such as the use of further inspections, the issuance of a Director's Order (which would have required the WPUC to comply with the requirements for treatment and monitoring), or enforcement proceedings. The lack of any follow-up or enforcement efforts may have led the

Koebel brothers to believe the recommendations were not very important, even to the MOE (*flawed mental model*).

The WPUC Commissioners received a copy of the 1998 inspection report but did nothing beyond asking for an explanation from Stan Koebel and accepting his word that he would correct the deficient practices (*inadequate control*). They never followed up to make sure he did (*missing feedback*).

The mayor of Walkerton and the municipality also received the report but they assumed the WPUC would take care of the problems. When the local Walkerton public health inspector read the report, he filed it, assuming that the MOE would ensure that the problems identified were properly addressed. Note the *coordination problems* here in an area of *overlapping control*. Both the MOE and the local public health inspector should have followed up on the 1998 inspection report, but there was no written protocol instructing the public health inspector on how to respond to adverse water quality reports or inspection reports. The MOE also lacked such protocols. The Province's water safety control structure had clearly become ineffective.

A final important change in the safety control structure involved the drinking water surveillance program in which the MOE monitored drinking water across the province. In 1996, the Provincial government dropped E. coli testing from its Drinking Water Surveillance Program. The next year, the Drinking Water Surveillance Program was shut down entirely (*feedback loop eliminated*). At the same time, the provincial government directed MOE staff not to enforce dozens of environmental laws and regulations still on the books (*control algorithms eliminated*). Farm operators, in particular, were to be treated with understanding if they were discovered to be in violation of livestock and waste-water regulations. By June, 1998, the Walkerton town council was concerned enough about the situation to send a letter directly to the Premier (Mike Harris), appealing for the province to resume testing of municipal water. There was no reply.

MOE officials warned the government that closing the water testing program would endanger public health. Their concerns were dismissed. In 1997, senior MOE officials drafted another memo that the government *did* heed. This memo warned that cutbacks had impaired the Ministry's ability to enforce environmental regulations to the point that the Ministry could be exposed to lawsuits for negligence if and when an environmental accident occurred. In response, the Provincial government called a meeting of the Ministry staff to discuss how to protect itself from liability, and it passed a Bill ("The Environmental Approvals Improvement Act") that, among other things, prohibited legal action against the government by anyone adversely affected by the Environment Minister's failure to apply environmental regulations and guidelines.

Many other groups warned senior government officials, ministers, and the Cabinet of the danger of what it was doing, such as reducing inspections and not making the notification guidelines into regulations. The warnings were ignored. Environmental groups prepared briefs. The Provincial Auditor, in his annual reports, criticized the MOE for deficient monitoring of groundwater resources and for failing to audit small water plants across the province. The International Joint Commission expressed its concerns about Ontario's neglect of water quality issues, and the Environmental Commissioner of Ontario warned that the government was compromising environmental protection, pointing specifically to the testing of drinking water as an area of concern.

In January 2000 (three months before the Walkerton accident), staff at the MOE's Water Policy Branch submitted a report to the Provincial government warning that "Not monitoring drinking water quality is a serious concern for the Ministry in view of its mandate to protect public health."

The report stated that a number of smaller municipalities were not up to the job of monitoring the quality of their drinking water. It further warned that because of the privatization of the testing labs, there was no longer a mechanism to ensure that the MOE and the local Medical Officer of Health were informed if problems were detected in local water systems. The Provincial government ignored the report.

The warnings were not limited to groups or individuals. Many adverse water quality reports had been received from Walkerton between 1995 and 1998. During the mid to late 1990s, there were clear indications that the water quality was deteriorating. In 1996, for example, hundreds of people in Collingswood (a town near Walkerton) became ill after cryptosporidium (a parasite linked to animal feces) contaminated the drinking water. Nobody died, but it should have acted as a warning that the water safety control structure had degraded. Between January and April of 2000 (the months just prior to the May E. coli outbreak), the lab that tested Walkerton's water repeatedly detected coliform bacteria—an indication that surface water was getting into the water supply. The lab notified the MOE on five separate occasions. The MOE in turn phoned the WPUC, was assured the problems were being fixed, and let it go at that (*inadequate control*). The MOE failed to inform the Medical Officer of Health, as by law it was required to do (*communication flaw*). One of the reasons for the delay in issuing a boil water advisory when the symptoms of E. coli contamination started to appear in Walkerton was that the latest report in the local Health Unit's files of any problems with the water was over two years old (*incorrect mental model*). In May 2000, Walkerton changed its testing lab to A&L Canada who, as noted above, did not know about the reporting guidelines.

The Walkerton Inquiry report notes that the decisions to remove the water safety controls in Ontario or to reduce their enforcement were taken without an assessment of the risks or the preparation of a risk management plan. The report says there was evidence that those at the most senior levels of government who were responsible for the decisions considered the risks to be manageable, but there was no evidence that the specific risks were properly assessed or addressed.

All of these changes in the Ontario water safety control structure over time led to the modified control structure shown in Figure 2. One thing to notice in comparing Figure1 and Figure 2 is the disappearance of many of the feedback loops. When the models are shown on a computer, graphics can be used to illustrate and assist in understanding the changes in the control structure over time.

Dynamic Process Model

As we have seen, the system's defenses or safety controls may degrade over time due to changes in the behavior of the components of the safety control loop. The reasons for the migration of the system toward a state of higher risk will be system specific and can be quite complex. In contrast to the usually simple and direct relationships represented in event chain accident models, most accidents in complex systems involve relationships between events and human actions that are highly non-linear, involving multiple feedback loops. The analysis or prevention of these accidents therefore requires an understanding not only of the static structure of the system (the *structural complexity*) and of the changes to this structure over time (the structural *dynamics*), but also the dynamics behind these changes (the *behavioral* or *dynamic complexity*). The previous section presented an approach to describing and analyzing the static safety control structure and how to use that to describe the changes to that structure that occur over time. This section presents a way to model and understand the dynamic processes behind the changes to the static

control structure and *why* it changed over time, potentially leading to ineffective controls and unsafe or hazardous states.

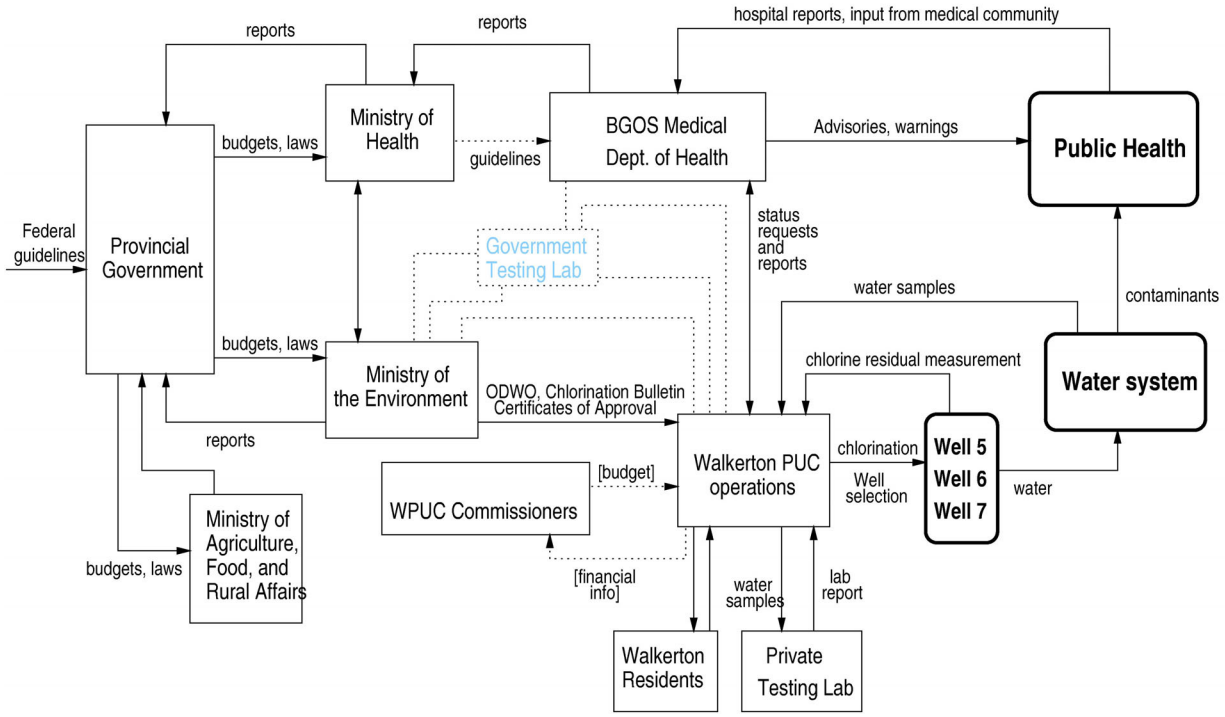


Figure 2 - The Basic Water Safety Control Structure at the Time of the Accident. Dotted lines represent communication, control, or feedback channels that had disappeared or become ineffective.

The approach proposed uses the modeling techniques of system dynamics. The field of system dynamics, created at MIT in the 1950's by Jay Forrester, is designed to help decision makers learn about the structure and dynamics of complex systems, to design high leverage policies for sustained improvement, and to catalyze successful implementation and change. Drawing on engineering control theory and the modern theory of nonlinear dynamical systems, system dynamics involves the development of formal models and simulators to capture complex dynamics and to create an environment for organizational learning and policy design.

These ideas are particularly relevant when analyzing system accidents. The world is dynamic, evolving, and interconnected, but we tend to make decisions using mental models that are static, narrow, and reductionist. Decisions that might appear to have no effect on safety—or even appear to be beneficial—may in fact degrade safety and increase risk. Using system dynamics, one can, for example, understand and predict instances of policy resistance or the tendency for well-intentioned interventions to be defeated by the response of the system to the intervention itself. A companion paper submitted to this workshop presents archetypical system dynamic models often associated with accidents.

Figure 3 shows a system dynamics model for the Walkerton accident. The basic structures in the model are variables, stocks (represented by rectangles), and flows (double arrows into and out of stocks). Lines with arrows between the structures represent causality links, with a positive

polarity meaning that a change in the original variable leads to a change in the same direction in the target variable. Similarly, a negative polarity means that a change in the original variable leads to a change in the opposite direction of the target variable. Double lines across a link represent a delay. Delays introduce the potential for instabilities in the system.

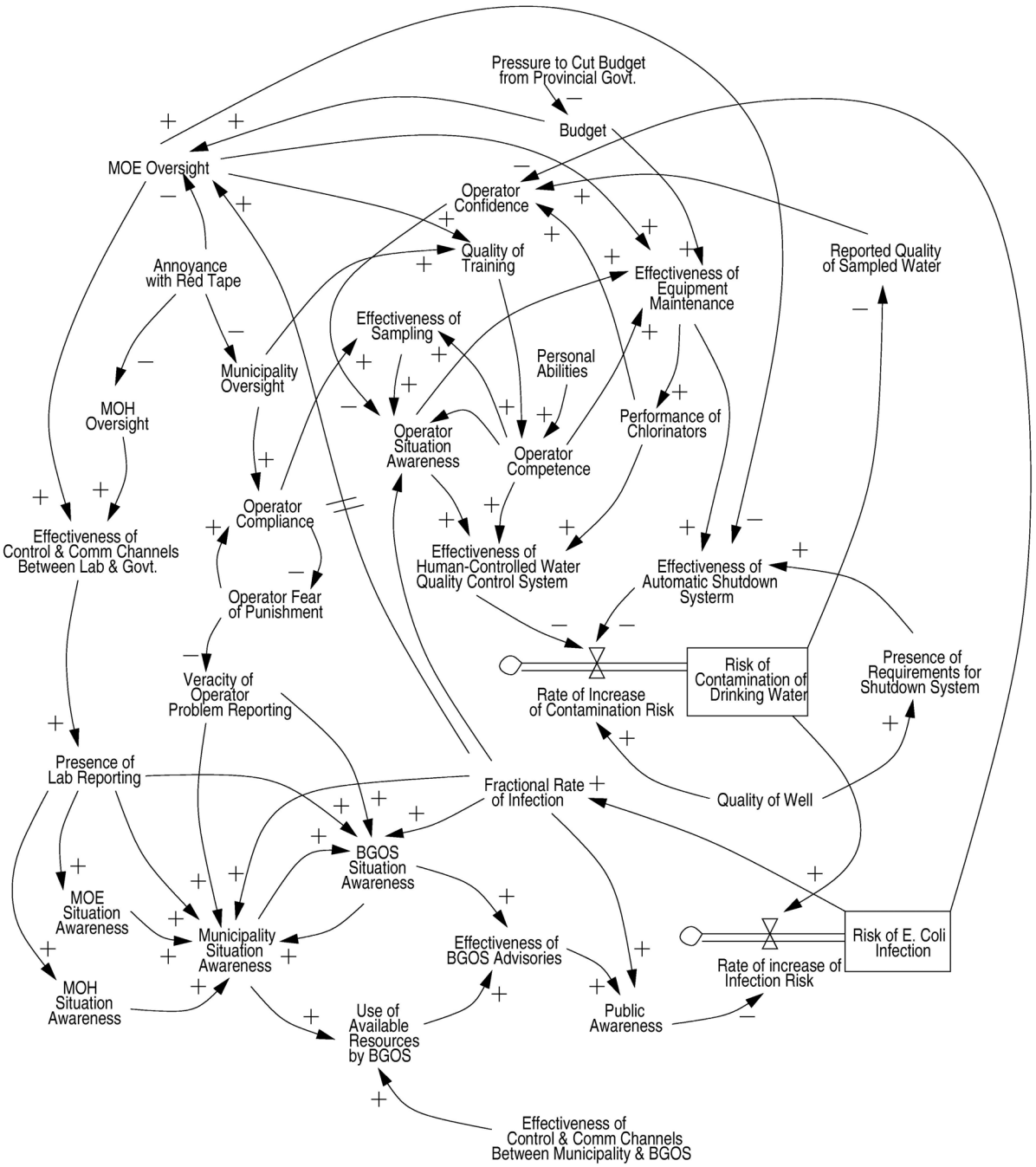


Figure 3 - A Systems Dynamics Model for the Walkerton Water Contamination Accident

Modeling the entire system dynamics is usually impractical. The challenge is to choose relevant subsystems and model them appropriately for the intended purpose. STAMP provides the guidance for determining what to model when the goal is risk management. In the example provided, we focused primarily on the organizational factors, excluding the physical processes allowing the mixing of manure with the source water. Depending on the scope or purpose of the model, different processes could be added or removed.

In complex systems, all dynamics, despite their complexity, arise from two types of feedback loops [9]: positive (reinforcing) and negative (balancing). In system dynamics terms, degradation over time of the safety control structure, as represented by reinforcing loops, would lead inevitably to an accident, but there are balancing loops, such as regulation and oversight that control those changes. In Ontario, as feedback and monitoring controls were reduced, the mental model of the central government leaders and the ministries responsible for water quality about the current state of the water system became increasingly divorced from reality. A belief that the water quality controls were in better shape than they actually were led to disregarding warnings and continued reduction in what were regarded as unnecessary regulation and red tape.

Accidents occur when the balancing loops do not adequately overcome the influences degrading the safety controls. Understanding why this degradation occurred (why risk increased) is an important part of understanding why the accident occurred and learning how to prevent repetitions in the future, i.e. how to set up more effective safety control structures. It is also an important part of identifying when the socio-technical system is moving toward a state of unacceptable risk.

Our Walkerton model includes a number of exogenous variables (pressure to reduce the size of government and cut budgets, attempts to reduce business and government red tape, etc.) that act as levers on the behaviors of the system. When these variables are changed without any consideration of the dynamics of the system being modeled, the effectiveness of the safety control structure can deteriorate progressively, with few if any visible signs. For instance, the attempts to reduce red tape decreased the oversight of the ministries and municipalities. This decrease in oversight in turn had a negative effect on the control and communication channels between the government and the laboratories performing water quality analyses. Eventually, the laboratories stopped reporting the results of the tests. Because of this lack of reporting, the Walkerton municipality was much slower to realize that the water was contaminated, leading to a delay in the mobilization of the resources needed to deal with the contamination, and the effectiveness of the advisories issued was thus greatly diminished, increasing the risk of infection in the population.

Accident investigations often end with blame being assigned to particular individuals, often influenced by legal or political factors. The system dynamics models, on the other hand, can show how the attitude and behavior of individuals is greatly affected by the rest of the system and how and why such behavior may change over time. For instance, operator competence depends on the quality of training, which increases with government oversight but may decrease over time without such oversight due to competing pressures. An operator's fear of punishment, which in this case led Stan Koebel to lie about the adverse water quality test reports, is balanced by compliance with existing rules and regulations. This compliance, in turn, is directly influenced by the extent of government oversight and by the government's response to similar behavior in the past.

Note that even though the STAMP analysis of the Walkerton water system contamination provided thus far has not yet even gotten to the point where most accident investigations start—

the proximate events to the loss—it is clear that the system was in a state where the risk of an accident was very high and a lot of different scenarios (or triggers) could have led to a tragedy. Most of the information required to understand the reasons for this accident (or at least the context in which it happened and why it was likely to occur) are outside the usual proximate chain of events used to describe the cause of an accident and to identify a "root cause."

Summary Accident Analysis (Causal Analysis)

At this point it is now possible to show the proximate events and see how they combined with the inadequate safety control structure at the time to lead to the Walkerton E. coli contamination. A STAMP analysis interprets the events not in terms of a causal chain but in terms of the implications and feedback relationships on the safety control structure. For space reasons, we will not repeat a description of the events nor show them on the control structure.

The final model, the summary accident analysis, consists of a description of the inadequate control actions by each of the components in the water safety control structure and the reasons for these actions using the accident factors in STAMP (e.g., flawed mental models, lack of coordination among controllers, inadequate control algorithms or inadequate execution of acceptable control algorithms, missing feedback loops, etc.). The Appendix contains the final accident analysis model for the Walkerton accident.

The final accident summary, along with the systems dynamics model, provides the information necessary for devising recommendations and changes that do not simply fix symptoms but eliminate the root causes (the inadequate control structure) of the accident. Despite the government's argument that the accident was solely due to actions by Stan Koebel and the WPUC, after the accident many recommendations and changes were made to fix the problems noted here including establishing regulatory requirements for agricultural activities with potential impacts on drinking water sources, updating of standards and technology, improving current practices in setting standards, establishing legally enforceable regulations rather than guidelines, requiring mandatory training for all water system operators and requiring grandfathered operators to pass certification examinations within two years, developing a curriculum for operator training and mandatory training requirements specifically emphasizing water quality and safety issues, adopting a province-wide drinking water policy and a Safe Drinking Water Act, strictly enforcing drinking water regulations, and committing sufficient resources (financial and otherwise) to enable the MOE to play their role effectively.

Conclusions and Future Work

The Walkerton Inquiry report did an excellent job, which is why the information was available to create the STAMP models. Most accident reports do not dig as deeply into the root causes of the accident. STAMP was developed to assist in determining what questions should be asked during investigations to maximize the learning process.

The use of a systems-theoretic accident model like STAMP does not lead to identifying single causal factors or variables. It will thus not be terribly satisfying to those focused on finding someone or something to blame. It does, however, a much better job than chain of events models in providing information about the changes that are needed to prevent accidents in the future, particularly changes to the organizational structure and to engineering design, manufacturing, and operations.

Our future goals are to add more sophisticated human error and decision making models to STAMP, to apply the model to hazard analysis and accident prevention, and to explore the implications for new approaches to risk assessment and risk management. We are also working on tool support for graphically displaying and animating the models (including simulating the system dynamics models) and for providing assistance in creating them.

References

1. Leveson, Nancy G. (1995) *Safeware: System Safety and Computers*, Addison-Wesley.

2. Leveson, Nancy G. (2003) A New Accident Model for Engineering Safer Systems, to appear in *Safety Science*, Elsevier Science Ltd.

3. Rasmussen, Jens (1997) Risk Management in a Dynamic Society: A Modeling Problem, *Safety Science*, Vol. 27, No. 2/3, Elsevier Science Ltd., pp. 183–213.

4. O'Connor, Dennis R. (2002) Report of the Walkerton Inquiry, Ontario Ministry of the Attorney General.

5. Leplat, Jacques (1987) Occupational Accident Research and Systems Approach, in Jens Rasmussen, Keith Duncan, and Jacques Leplat (eds.) New Technology and Human Error, John Wiley & Sons, New York, pp. 181–191.

6. Checkland, Peter (1981) Systems *Thinking, Systems Practice*, John Wiley & Sons, New York.

7. Vicente, Kim and Christoffersen, K. (2003) The Walkerton E. coli Outbreak: A Test of Rasmussen's Framework for Risk Management in a Dynamic Society, *Theoretical Issues in Ergonomics Science*, in press.

8. Diemer, Ulli (2000) The Poisonous Legacy of Ontario's Environment Cutbacks, *Canada Dimension Magazine*, July-August

9. Sterman, John (2000) *Business Dynamics*, Mc-Graw Hill.

Biographies

Nancy Leveson, MIT; Cambridge, MA, U.S.A.; telephone +1.617.258.0505; leveson@mit.edu; Dr. Leveson is Professor of Aeronautics and Astronautics and also Professor of Engineering Systems at MIT.

Mirna Daouk, MIT; Cambridge, MA, U.S.A.; daouk@mit.edu; Ms. Daouk is a Ph.D. student at MIT in the Department of Aeronautics and Astronautics.

Nicolas Dulac, MIT; Cambridge, MA, U.S.A.; telephone +1.617.258.0505; ndulac@mit.edu; Mr. Dulac is a Ph.D. student at MIT in the Department of Aeronautics and Astronautics

Karen Marais, MIT; Cambridge, MA, U.S.A.; telephone +1.617.258.0505; karenm@mit.edu; Ms. Marais is a doctoral candidate at MIT in the Department of Aeronautics and Astronautics.

APPENDIX - Summary of Accident Causal Factors

Provincial Government

Safety Requirements and Constraints:

- Establish regulatory bodies and codes of responsibilities, authority, and accountability for the province.
- Provide adequate resources to regulatory bodies to carry out their responsibilities.
- Provide oversight and feedback loops to ensure that provincial regulatory bodies are doing their job adequately.
- Ensure adequate risk assessment is conducted and effective risk management plan is in place.
- Enact legislation to protect water quality.

Context in Which Decisions Made:

- Anti-regulatory culture.
- Efforts to reduce red tape.

Inadequate Control Actions:

- No risk assessment or risk management plan created to determine extent of known risks, whether risks should be assumed, and if assumed, whether they could be managed.
- Privatized laboratory testing of drinking water without requiring labs to notify MOE and health authorities of adverse test results. (Privatizing without establishing adequate governmental oversight)
- Relied on guidelines rather than legally enforceable regulations.
- No regulatory requirements for agricultural activities that create impacts on drinking water sources.
- Spreading of manure exempted from EPA requirements for Certificates of Approval
- Water Sewage Services Improvement Act ended provincial Drinking Water Surveillance program
- No accreditation of water testing labs (no criteria established to govern quality of testing personnel, no provisions for licensing, inspection, or auditing by government).
- Disbanded ACES.
- Ignored warnings about deteriorating water quality.
- No law to legislate requirements for drinking water standards, reporting requirements, and infrastructure funding.
- Environmental controls systematically removed or negated.

Feedback:

- No monitoring or feedback channels established to evaluate impact of changes

Ministry of the Environment

Safety Requirements and Constraints:

- Ensure those in charge of water supplies are competent to carry out their responsibilities.
- Perform inspections and enforce compliance if problems found.
- Perform hazard analyses to provide information about where vulnerabilities are and monitor them.
- Perform continual risk evaluation of existing facilities and establish new controls if necessary.
- Establish criteria for determining whether a well is at risk.
- Establish feedback channels for adverse test results. Provide multiple paths so that dysfunctional paths cannot prevent reporting.
- Enforce legislation, regulations, and policies applying to construction and operation of municipal water systems.
- Establish certification and training requirements for water system operators.

Context in Which Decisions Made:

- Critical information about history of known vulnerable water sources not easily accessible.
- Budget cuts and staff reductions

Inadequate Control Actions:

- No legally enforceable measures taken to ensure that concerns identified in inspections are addressed. Weak response to repeated violations uncovered in periodic inspections.
- Relied on voluntary compliance with regulations and guidelines.
- No systematic review of existing certificates of approval to determine if conditions should be added for continuous monitoring.
- Did not retroactively apply new approvals program to older facilities when procedures changed in 1992.
- Did not require continuous monitoring of existing facilities when ODWO amended in 1994.
- MOE inspectors not directed to assess existing wells during inspections.
- MOE inspectors not provided with criteria for determining whether a given well was at risk. Not directed to examine daily operating sheets.
- Inadequate inspections and improperly structured and administered inspection program..
- Approval of Well 5 without attaching operating conditions or special monitoring or inspection requirements.
- No followup on inspection reports noting serious deficiencies.
- Did not inform Walkerton Medical Officer of Health about adverse test results in January to April 2000 as required to do.
- Private labs not informed about reporting guidelines.
- No certification or training requirements for grandfathered operators.
- No enforcement of continuing training requirements.
- Inadequate training of MOE personnel.

Mental Model Flaws:

- Incorrect model of state of compliance with water quality regulations and guidelines.
- Several local MOE personnel did not know E. coli could be fatal.

Feedback:

- Did not monitor effects of privatization on reporting of adverse test results.

Coordination:

- Neither MOE nor MOH took responsibility for enacting notification legislation.

Ministry of Health

Safety Requirements and Constraints:

- Ensure adequate procedures exist for notificatoin and risk abatement if water quality is compromised.

Inadequate Control Actions:

- No written protocol provided to local public health inspector on how to respond to adverse water quality or inspection reports.

Local (BGOS) Medical Dept. of Health

Safety Requirements and Constraints:

- Provide oversight of drinking water quality.
- Follow up on adverse drinking water quality reports.
- Issue boil water and other advisories if public health at risk.

Context in Which Decisions Made:

- Most recent water quality reports over 2 years old.
- Illness surfacing in communities outside Walkerton
- E. coli most commonly spread through meat.

Inadequate Control Actions:

- Advisory delayed.
- Advisory should have been more widely disseminated.
- Public health inspector did not follow up on 1998 Walkerton inspection report.

Mental Model Flaws:

- Thought were receiving adverse water quality reports after privatization.
- Unaware of reports of E. coli linked to treated water source.
- Thought Stan Koebel was relaying the truth.

Coordination:

- Assumed MOE was ensuring inspection report problems were resolved.

A&L Canada Laboratories

Safety Requirements and Constraints:

- Provide timely and accurate reports on testing results to MOE, WPUC, and Medical Dept. of Health (MOH)

Inadequate Control Actions:

- Did not follow provincial guidelines and inform MOE and MOH of adverse test results.

Mental Model Flaws:

- Did not know about reporting guidelines;
- Considered results to be proprietary.

WPUC Commissioners

Safety Requirements and Constraints:

- Oversee operations to ensure water quality is not compromised.

Context in Which Decisions Made:

- Elected officials
- No training or educational requirements.

Inadequate Control Actions:

- Relied on Stan Koebel to identify and resolve any concerns related to operation of the water supply. Did not monitor to ensure problems fixed.
- Did not establish, oversee, nor enforce policies and practices for operating the Walkerton public water system.
- Concentrated only on financial matters.

Mental Model Flaws:

- Little knowledge of water safety and operation of system.;
- Unaware of improper treatment and monitoring practices of WPUC operators.

Walkerton PUC Operations Management

Safety Requirements and Constraints:

- Monitor operations to ensure that sample taking and reporting is accurate and adequate chlorination is being performed.
- Keep accurate records.
- Update knowledge as required.

Context in Which Decisions Made:

- Complaints by citizens about chlorine taste in drinking water.
- Improper activities were established practice for 20 years.
- Lacked adequate training and expertise.

Inadequate Control Actions:

- Inadequate monitoring and supervision of operations
- Adverse test results not reported when asked.
- Problems discovered during inspections not rectified.
- Inadequate response after first symptoms in community
- Did not maintain proper training or operations records.

Mental Model Flaws:

- Believed sources for water system were generally safe. Thought untreated water safe to drink.
- Did not understand health risks posed by underchlorinated water.
- Did not understand risks of bacterial contaminants like E. coli.
- Did not believe guidelines were a high priority.

Local Operations

Safety Requirements and Constraints:

- Apply adequate doses of chlorine to kill bacteria.
- Measure chlorine residuals.

Context in Which Decisions Made:

- Lacked adequate training.

Inadequate Control Actions:

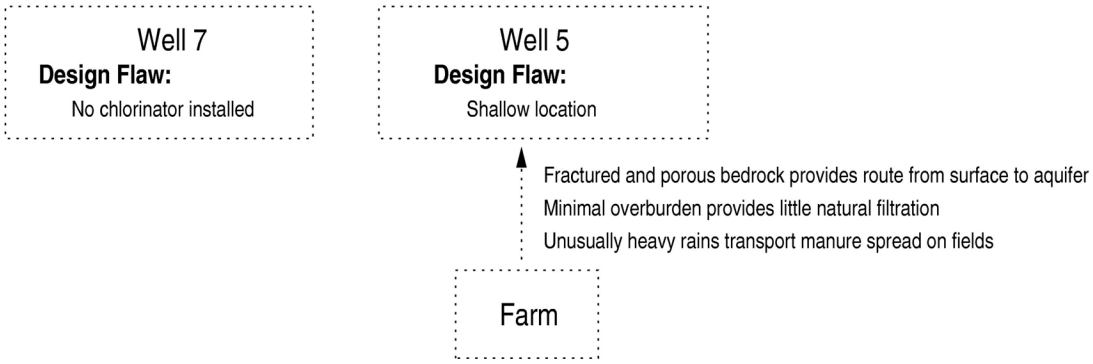
- Did not measure chlorine residuals on most days. Only started measuring in 1998.
  Made fictitious entires for residuals in daily operating sheets.
- Misstated locations from which samples had been collected.
- Did not use adequate doses of chlorine.
- Did not take measurements of chlorine residuals for Well 5 between May 13 and May 15 (after symptoms of problems appeared).
- Operated Well 7 without a chllorinator.

Mental Model Flaws:

- Inadequate training led to inadequate understanding of job responsibilities.
- Thought convenience was acceptable basis for sampling.

Physical Process

Context: Greatly increased farm operations.



Causal Determination in Road Accidents: An Application of the Halpern/Pearl Notion of' 'Actual Cause'

Gary A. Davis, University of Minnesota, Minneapolis, MN, U.S.A.
Tait Swenson, URS Corporation, Minneapolis, MN, U.S.A.

Keywords: road accidents, causation, counterfactuals

Abstract

Determining whether or not an event was a cause of a road accident often involves determining the truth of a counterfactual conditional, in which what happened is compared to what would have happened had the putative cause been absent. Using structural equation models, Pearl and his associates have recently developed a rigorous method for posing and answering causal questions, and this approach is especially well-suited to the analysis of road accidents. Following a general discussion of causal analysis, we apply these methods to a freeway rear-end collision. The results suggest that not only were the actions of the drivers actually involved in the collision causes of the accident, but so were the actions of drivers ahead of them.

Introduction

Although the costs and consequences of any particular road accident rarely approach those that occur in aircraft or rail accidents, the sheer number of road accidents occurring in a given year means that their total costs usually outstrip those from accidents in other modes. A road accident may be investigated by the police, in order to assess the possibility of criminal liability, by an accident investigator retained by a party involved in civil proceedings, by a governmental agency seeking to identify actions which could prevent similar accidents in the future, or by researchers seeking to advance our understanding of how and why accidents occur. All these investigative activities share a common concern however, to identify those events that could be considered as causes of the accident. For example, the Uniform Vehicle Code [10] states that to be guilty of vehicular homicide a driver must have been "engaged in the violation of any state law or municipal ordinance," and that "such violation is the proximate cause of said death." In tort law, "The most basic element of any tort cause of action is some causal connection between the act or omission of the tortfeasor and the plaintiff's injury" [6]. A main objective of an investigation by the National Transportation Safety Board (NTSB) is a statement of the "probable causes" of an accident, while the objective of the Tri-Level study was to provide "up to date data regarding traffic accident causation" [12].

Causal Concepts

Baker [1] has noted that causal attributions in road safety take a number of forms, and are often invoked to achieve rhetorical, rather than scientific, objectives. He has also given an often-used definition of "causal factor" as a circumstance "contributing to a result without which the result could not have occurred." This definition is (not by accident) similar to definitions of cause used in other types of accident investigation. For instance, Miller [9] points out that in a definition used by the NTSB, a "condition or event" qualifies as a probable cause of an accident if "had the condition or event been prevented…the accident would not occur," while the Air Force has used "A cause is an act, omission, condition, or circumstance which if corrected, eliminated, or avoided would have prevented the mishap." These definitions in turn share content with the legal notion of "cause in fact," where for an event to be considered as a cause it must satisfy a "but for" test, that is, "defendant's conduct is not a cause of the event, if the event would have occurred without it" [6].

Implicit in these ideas is first, that removal of a cause should be sufficient to prevent the result, and second that one determines whether or not a circumstance is a cause by carrying out a counterfactual test, where what happened is compared to what would have happened had the circumstance been absent. In practice however giving a rigorous yet general specification for such tests has proved somewhat daunting, the main challenge being to unambiguously specify what should count as the counterfactual condition. Since one can, with sufficient imagination, almost always describe a number of different scenarios where an accident is avoided, this test condition should involve a change that is in some sense minimal. Lewis [8] has given a philosophical treatment of truth conditions for causal assertions using a comparison between what actually happened and what happens in a closest possible world where certain counterfactual assertions are true. What is meant by "closest possible world" is left deliberately vague, which improves the generality of Lewis' treatment but makes it difficult to apply to practical cases. Over the past 15 years or so however, there has been increased interest in causal inference as a component of artificial intelligence, and one especially useful approach is based on what Pearl [11] calls a "causal model." This consists of a set of exogenous variables, a set of endogenous variables, and for each endogenous variable a structural equation describing how that variable changes in response to changes in the exogenous and/or other endogenous variables. Events are defined in terms of values taken on by the model's variables, and the closest possible world where a set of variables takes on (counterfactual) values can be unambiguously defined as the outcome of a modified causal model, where the exogenous variables are set to the same values as in the actual condition, but where the structural equations associated with the counterfactual event are replaced by assignment statements. Pearl goes on to describe how when the evidence about an event is not sufficient to uniquely identify the values taken on by each exogenous variable (i.e. to identify which possible world is the actual world) uncertainty can be accommodated by first placing a prior probability distribution over the causal model's exogenous variables and then using Bayesian updating to compute the posterior distribution given the evidence at hand. The probability attached to an assertion, either indicative or counterfactual, is then simply the posterior probability assigned to the set of possible worlds where that assertion is true. More recently, Halpern and Pearl [5] have extended these ideas and defined an "actual cause" as an event satisfying a "but for" test along with additional conditions which deal with some counterintuitive consequences of simple "but for" tests.

Application to Freeway Rear-ending Collisions

Over the past several years we have been applying Pearl's approach to the analysis of road accidents, with an emphasis on determining the degree to which excess speed could be considered a causal factor. Descriptions of some of this work can be found in [3,4]. In this paper though we would like to describe some preliminary results from an ongoing study of freeway rear-ending accidents. Although such accidents do not usually result in fatal or even very severe injuries, they are responsible for a substantial fraction of the unpredictable delays many of us now regard as unavoidable aspects of urban life. Frequently, such accidents occur when a platoon of vehicles successively brake and the braking deceleration of at least one vehicle is not sufficient to prevent it from colliding with the vehicle ahead. Reducing the frequency of such collisions, for example by improving the competency of drivers or deploying in-vehicle collision avoidance technology, could then be one way to reduce travel delay without resorting to expensive additions to highway capacity. In Minnesota, as is many other places, it is recommended that drivers maintain following headways of at least 2.0 seconds, and responsibility for a rear-end collision is generally attributed to the following vehicle actually involved in the collision. If however the actions of drivers earlier in the sequence also contribute to the collision, this method of giving feedback will leave these earlier drivers unaware of their contribution, and so be of limited effectiveness. But how can we assess the causal contributions, if any, of these other drivers?

These concerns are not new, and Brill [2] has described a relatively simple kinematic model of successive braking which applies to the problem at hand. Imagine a platoon of vehicles, indexed in order from first to last, by k=1,..,n, and let $v_1,v_2,...v_n$ denote their speeds. At time t=0 the lead driver brakes to a stop, with deceleration a_1, and after a reaction time r_2 driver 2 also brakes to stop, with deceleration a_2, and so forth. A rear-end collision between vehicles k and k+1 will be avoided as long as the distance needed by driver k+1 to stop does not exceed the available stopping distance. That is,

$$x_{k+1} + \frac{v_k^2}{2a_k} \geq r_{k+1}v_{k+1} + \frac{v_{k+1}^2}{2a_{k+1}} \qquad (1)$$

where x_{k+1} is the distance separating vehicle k's rear bumper from vehicle k+1's front bumper. Letting $x_{k+1}=v_{k+1}h_{k+1}$ express this distance in terms of driver k+1's speed and following headway, driver k+1 will stop before colliding if his or her deceleration satisfies

$$a_{k+1} \geq \frac{v_{k+1}^2}{\frac{v_k^2}{a_k} + 2v_{k+1}(h_{k+1} - r_{k+1})} \qquad (2)$$

Relation 2 has some interesting implications. Other things equal, the minimum deceleration required of driver k+1 increases as the deceleration used by driver k increases, since k+1's available stopping distance decreases as a_k increases. Also, other things equal, the minimum deceleration required by driver k+1 increases as the difference between k+1's following headway and reaction time (h_{k+1}-r_{k+1}) decreases. Together these features imply, as Brill pointed out, that if each driver in the platoon is a little slow in reacting, so that his or her reaction time is longer than the following headway, the minimum required deceleration will tend to increase for each succeeding vehicle. If the platoon is long enough a collision can become inevitable. In this case, it would appear reasonable to attribute the accident to the actions of each driver in the platoon, rather than to an egregious lapse by the last driver.

To illustrate how Halpern and Pearl's notion of actual cause might be applied to a freeway rear-end crash consider Figure 1, which displays Brill's sequential braking model (in this case involving a three-vehicle platoon) as a directed acyclic graph. The nodes of the graph represent the model's variables while the arrows indicate the presence and direction of causal dependencies. Those nodes without arrows pointing toward them (such as v_1) represent exogenous variables, while the others (such as a_{20}) represent endogenous variables. To complete the model we need to specify, for each endogenous variable, a structural equation. The variables a_{20} and a_{30} are the minimal decelerations needed, for vehicles 2 and 3 respectively, to stop before colliding with the vehicle ahead, and these are determined from the right-hand side of relation 2. We assume that the actual decelerations are then determined as

$$a_k = \min(a_{k0} + u_k, a) \qquad (3)$$

where a is a maximum achievable deceleration, and u_k is an exogenous term which accounts for the difference between observed and minimum deceleration. Finally, the variable y is a collision indicator, and is assumed to be determined via

$$y = \begin{cases} 0, \text{ if } a_{30} \leq a \\ 1, \text{ if } a_{30} > a. \end{cases} \qquad (4)$$

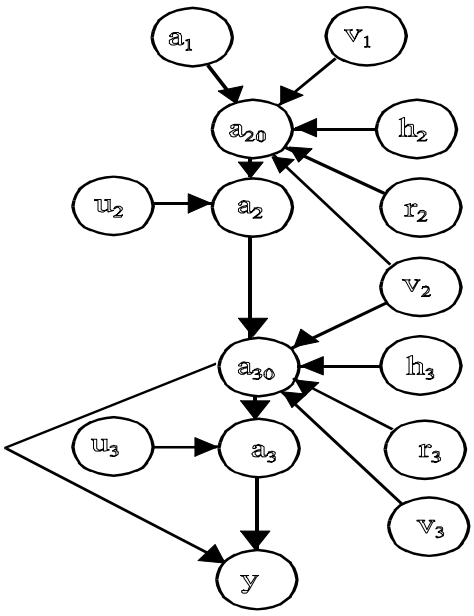


Figure 1 - Directed Acyclic Graph Representation of Three-Vehicle Platoon Collision Model.

For example, suppose $v_1=v_2=v_3=12.2$ meters/sec, that the maximum achievable deceleration is $a= 6.1$ meters/sec^2, and that driver 1 brakes to a stop with $a_1=1.5$ meters/sec^2. Suppose also that $h_2=2$ seconds but $r_2=4$ seconds, so that by relation 2 driver 2's minimum deceleration is $a_{20}=3.0$ meters/sec^2. Driver 2 then decelerates at 3.2 meters/sec^2 (which means that $u_2=0.2$ meters/sec^2), but suppose driver 3 is tailgating a bit, with $h_3=1.5$ seconds, and reacts after $r_3=2.5$ seconds. The minimum deceleration for driver 3 is then $a_{30}=6.7$ meters/sec^2, which exceeds the maximum deceleration $a=6.1$ meters/sec^2, and a rear-end collision between vehicle 2 and vehicle 3 occurs. Driver 3's tailgating can be considered an actual cause of this collision, since if we counterfactually set $h_2=2.0$ seconds but fix v_2, a_2 and v_3 at their actual values, the minimum deceleration needed by driver 3 falls to $a_{30}=4.3$ meters/sec^2, and the collision is avoided. But driver 2's long reaction time is also an actual cause of the collision, since setting $r_2=2.5$ seconds, but keeping $u_2=0.2$ meters/ sec, leads to $a_{30}=2.7$ meters/sec^2.

An Actual Collision

Do similar things happen in reality? As part of an ongoing study, permanently mounted video cameras were installed on high-rise buildings adjacent to an urban freeway in Minnesota. The cameras were connected to a computer that recorded the weekday traffic movements from the early morning rush hour to the early evening. Video records were saved in one-hour segments on the computer's hard drive. Accident reports filed with the State Patrol along with incident reports recorded by the Minnesota Dept. of Transportation's Traffic Management Center were then used to determine which video segments might contain accident footage.

The computer program VideoPoint was used to extract the screen coordinates of vehicles from a frame of the recorded video by clicking on a discernable point on the object of interest. The program then advances the movie one frame and the process is repeated, so by successively clicking on the same point of a vehicle's image it was possible to record the sequence of coordinates representing the vehicle's trajectory. Standard photogrammetry transformations were then used to convert the screen

coordinates to the corresponding real-world coordinates. Figure 2 shows the trajectories of a platoon of seven vehicles involved in sequential braking maneuvers, recorded during an afternoon peak period, where the seventh vehicle was observed colliding with the sixth.

To assess the possible causal contributions of the drivers in this platoon, it was first necessary to determine values for the individual speeds, decelerations, reaction times and following headways. During the time before a vehicle began braking it was assumed that the vehicle traveled at a constant speed, and visual examination of the position-time diagram shown in Figure 2 supports this assumption. Each vehicle speed was then determined by fitting a linear regression line to the initial portion of its trajectory data and determining the slope of this best fitting line.

When braking began, it was assumed that each driver decelerated with the intention to stop. It was also assumed that the deceleration was constant over the braking period. This allowed a fairly straightforward determination of the decelerations, headways, and reactions times from the trajectories shown in Figure 2. The motion of each vehicle can be described using a two-part relation, where the first part gives the vehicle's trajectory before braking and the second part describes the distance traveled during braking. That is,

$$z_k(t) = \begin{cases} v_k t, & t \le t_{0k} \\ v_k t - 0.5\, a_k (t - t_{0k})^2, & t > t_{0k}. \end{cases} \qquad (5)$$

where t_{0k} is the time at which driver k began braking. Determining t_{0k}, and the deceleration a_k was accomplished by minimizing the sum of squared errors between the measured position value and the position value estimated using equation 5.

The term 'space headway' is used to describe the distance between two successive vehicle front ends at the instant the leading vehicle begins braking. These values were determined from the Figure 2 trajectories using the braking times estimated as described above, as depicted in Figure 3. Space headways were then converted to separation distances by subtracting a value of 4.6 meters for the effective length of the vehicle, and these were in turn converted into separation headways (h_k) by dividing by the speed of the following vehicle. Finally, reaction times were defined as the difference in time between when the leading vehicle began to brake and the time when the following vehicle began to brake, and Figure 4 illustrates how these were determined from the vehicle trajectories. The results of the data extraction are displayed in Table 1.

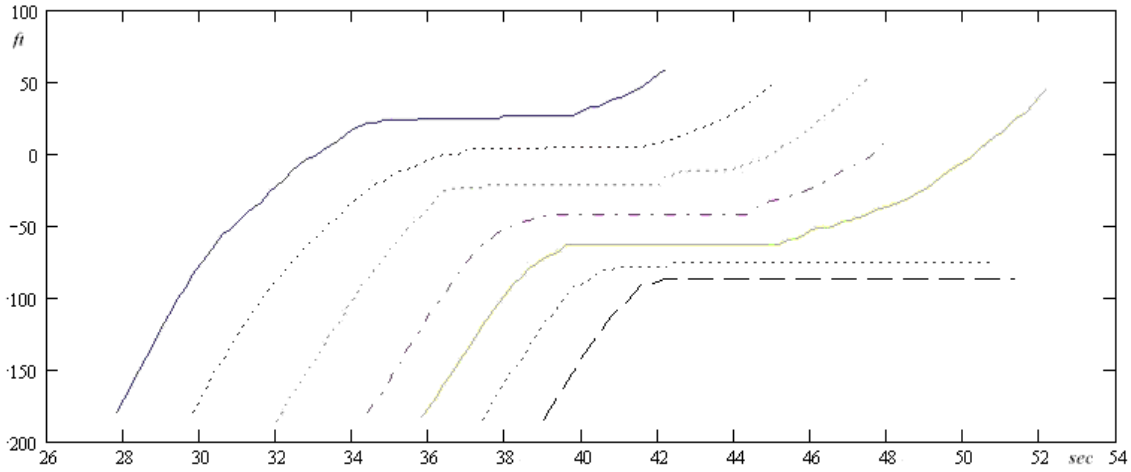


Figure 2 - Trajectories of Vehicles Involved in Actual Collision
x-axis is in seconds, y-axis is in feet.

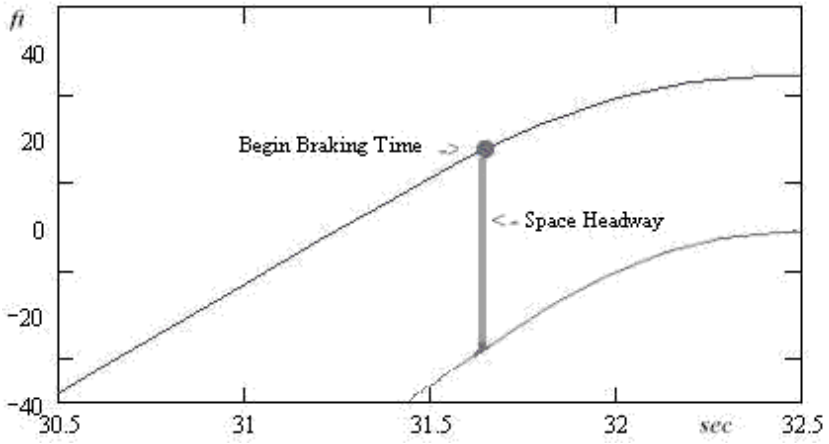


Figure 3 - Example of Space Headway Determination
x-axis is in seconds, y-axis is in feet.

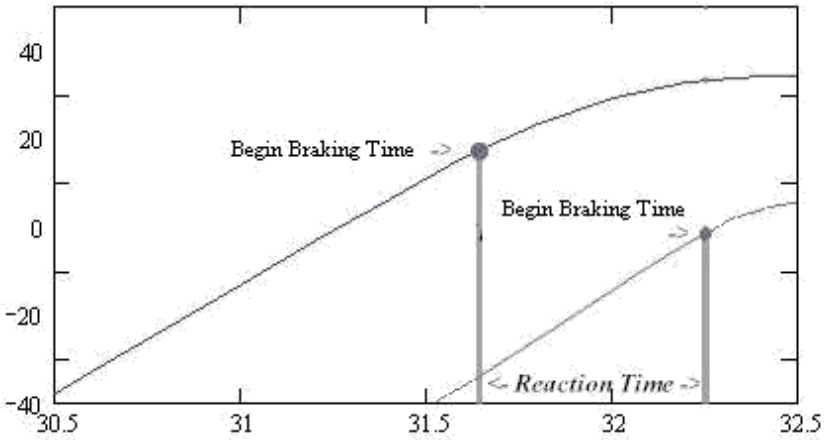


Figure 4 - Example of Reaction Time Determination
x-axis is in seconds, y-axis is in feet.

'Actual Cause' Analysis

The entries in Table 1 tell an interesting story. Driver 1 braked to a stop with a deceleration of 2.1 meters/sec^2, and about 1.5 seconds later driver 2 braked with a deceleration of about 1.8 meters/sec^2. It was possible for driver 2 to decelerate less rapidly than driver 1 because 2's following headway, at 1.7 seconds, was longer than 2's reaction time. Driver 3 on the other hand needed almost 4 seconds to react, and although 3's following headway was roughly equal to the recommended minimum of 2.0 seconds, 3's minimum deceleration jumped to about 3.7 meters/sec^2, with an actual deceleration of 4.1 meters/sec^2. Drivers 4 and 5 also had reaction times longer than their following headways, though the differences were not as extreme as 3's, so the minimum and actual braking decelerations continued to increase. Driver 6 was a bit more on the ball, but was traveling a bit faster than driver 5, and so the minimum deceleration increased again. When we come to driver 7, whose reaction time was about 0.5 seconds longer than his/her headway, the minimum deceleration jumped to 6.7 meters/sec^2, which exceeds the 5.8 meters/sec^2 observed to have been used by driver 7.

Table 1 - Values of Vehicle and Driver Variables for Seven-Vehicle Platoon.

	Variable Values					
Vehicle	**v_k (m/s)**	**h_k (sec)**	**r_k (sec)**	**a_k (m/s^2)**	**a_{k0} (m/s^2)**	**u_k (m/s^2)**
1	14.97	--	--	2.08	--	--
2	14.29	1.69	1.46	1.82	1.77	0.05
3	13.24	2.0	3.97	3.24	2.62	0.62
4	13.07	1.95	2.25	4.11	3.69	0.42
5	12.30	1.22	1.38	4.27	4.04	0.23
6	13.16	1.15	1.01	4.55	4.43	0.12
7	12.91	1.28	1.79	5.82	6.68	--

So who was responsible for this collision? Starting with driver 7, it is straightforward to verify that if 7 had had a following headway of 2.0 seconds, using the measured values for his/her speed and reaction time and driver 6's speed and deceleration, then 7's minimum deceleration drops to about 3.8 meters/sec^2. Since this is substantially lower than 7's observed deceleration of 5.8 meters/sec^2, we can conclude that driver 7's failure to observe the recommended following rule was an actual cause of the collision. But now let's look at driver 3. His or her reaction time was clearly long compared to what other drivers in the platoon appeared capable of, and we can ask whether or not this long reaction time might also have been an actual cause. Setting 3's reaction time to the counterfactual value of 2.5 seconds, leaving all observed speeds and headways, and all other observed reaction times alone, and then computing the actual decelerations for drivers 4, 5, and 6 by adding the observed differences u_k to the new computed minimums, produces a counterfactual minimum deceleration for driver 7 of about 3.9 meters/sec^2. Since this is also substantially lower than driver 7's observed deceleration, we can say that driver 3's long reaction time was also an actual cause. Next, looking at Table 1, we can see that both driver 4 and driver 5 also had observed reaction times that were longer than their headways, and we can ask whether or not these might also be considered causes of the collision. Separately setting the reaction times of drivers 3, 4 and 5 to their observed headways produced minimum decelerations for driver 7 of 3.5 meters/sec^2, 5.5 meters/sec^2, and 5.9 meters/sec^2, respectively. Finally, if drivers 3, 4,and 5 all had reaction times equal to their observed headways, the minimum deceleration for driver 7 would fall to 3.0 meters/sec^2.

Conclusion

It has been observed that at night drivers sometimes 'overdrive' their headlights, in that the stopping distances for their chosen speeds exceed the distances they can see ahead. The above results suggest that in congested conditions freeway drivers on occasion overdrive their reaction times, in the sense that their reaction times tend to be longer than their following headways. At least for this example, this over-driving appears to be locally benign, because based on what the vehicle ahead is doing and on an expectation that if the driver ahead does brake the deceleration will not be too extreme, then sufficient time to slow or stop is still available. What Brill's relation 2 shows though is that when each of a platoon of drivers overdrives their reaction times, this expectation of relatively gentle deceleration by the vehicle ahead can break down, so that in congested conditions prevention of rear-end collisions can require that drivers base their decisions on more than local information. Brill's effect can also be interpreted as resulting from the action of external costs. An over-driver will gain the benefits of his or her individual action (whatever those might be), while the costs of that action will tend to fall disproportionately on following drivers. This suggests that over-driving in congested conditions will be "consumed" at levels exceeding what is socially optimal. As with other situations involving external costs, achieving a socially optimal decision would then require some form of coordination mechanism.

More generally, Kletz [7] has argued that effective prevention of accidents not only requires identifying immediate causes, but also avoiding the accident by identifying those more distant causes that created the conditions making the accident possible. But determining whether or not an event qualifies as a cause requires a counterfactual test, and rational discussion can break down when different parties implicitly compare the actual to different "closest possible worlds." When the underlying mechanisms governing the accident process can be expressed as structural equations, Pearl has shown how to unambiguously define truth conditions for causal assertions, in a form that can be readily applied to actual cases.

Acknowledgements

This research was supported by the Intelligent Transportation Systems Institute at the University of Minnesota.

References

1. Baker, J.S. Causes and contributing factors in traffic accidents. in L. Fricke (ed.) *Traffic Accident Reconstruction*, Northwestern University Traffic Institute, 1990.
2. Brill, E. A car-following model relating reaction times and temporal headways to accident frequency. *Transportation Science*, 6, 1972, 343-353.
3. Davis, G. Towards a unified approach to causal analysis in traffic safety using structural causal models. In M. Taylor (ed.) *Transportation and Traffic Theory in the 21st Century*, Pergamon, 2002, 247-266.
4. Davis, G. and Pei, J. Bayesian networks and traffic accident reconstruction. *Proceedings of the 9th International Conference on Artificial Intelligence and Law*, Association for Computing Machinery, 2003, 171-176.
5. Halpern, J. and Pearl, J. Causes and explanations: A structural model approach-Part I: Causes. in J. Breese and D. Koller (eds.) *Uncertainty in Artificial Intelligence: Proceedings of the Seventeenth Conference*, Mortgan-Kaufman, 2001, 194-202.
6. Kionka, E. *Torts in a Nutshell, 2nd edition*. West Publishing Co., 1992.
7. Kletz, T. *Learning from Accidents, Third Edition*. Gulf Professional Publishing, 2001.
8. Lewis, D. Causation. *Journal of Philosophy*, 70, 1973, 556-567.
9. Miller, C. Aviation accident investigation: Functional and legal perspectives. *Journal of Air Law and Commerce*, 46, 1981, 237-293.
10. NCTLO. *Uniform Vehicle Code and Model Traffic Ordinance*. National Committee on Uniform Laws and Ordinances, 1992.
11. Pearl, J. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2000.
12. Treat, J. Tumbas, N., McDonald, S. Shinar, D. Hume, R. Mayer, R. Stansifer, R. and Castellon, N. *Tri-Level Study of the Causes of Traffic Accidents, Executive Summary*. National Highway Traffic Safety Administration, 1979.

Biography

Gary A. Davis, Associate Professor, Dept. of Civil Engineering, University of Minnesota, 500 Pillsbury Drive SE, Minneapolis, MN, 55455, U.S.A. telephone - +1.612.625.2598; fax- +1.612.626.7750; email- drtrips@umn.edu.

Tait Swenson, ITS Engineer, URS Corporation, 700 Third Street S, Minneapolis, MN, 55415, U.S.A. telephone-+1.612.370.0700; fax-+1.612.370.1378; email- tait_swenson@URSCorp.com.

Investigating Investigation Methodologies

Ludwig Benner Jr.; Starline Software Ltd.; Oakton Virginia U.S.A.

Keywords: investigation, methodology, RCA, MES

Abstract

Recent years have seen an increase in the number of investigation methodologies. For example, today investigators can choose among investigation methodologies like ICAO, ISIM. MES, Root Cause Analysis, Tripod-Beta, or TOR, to name a few. How do investigations and work products produced with these investigation methodologies compare?

Several investigators have reported comparisons of investigation methodologies but criteria for comparison, developed in different ways, have been inconsistent, leaving the question of their comparative merits unresolved. This paper is a progress report of an inquiry to develop a comparison based on a desktop simulation of an investigation with one methodology using data from a published accident report prepared with another methodology. The work is disclosing substantial differences attributable to methodologies.

Introduction

Choosing a methodology is a key investigation program decision. What information is available to support this decision? A recent paper by Sklet [1] presents a comparison of 14 different methods of investigation, offering seven different characteristics by which the methods are compared. Harvey [2] examined four accident investigation models in terms of their ability to satisfy five evaluation criteria representing his view of major purposes of accident investigation. The author [3] ranked 17 investigation models used by governmental organizations, using twelve assessment criteria. Henderson et al [4] reported five criteria for "good investigations." The Center for Chemical Process Safety [5] published a list of 18 investigation "techniques" with three basic criteria. The criteria for the these comparisons were all different.

Sklet developed seven characteristics for his comparison, derived from his expectations of the results required, and presented a table summarizing the attributes of each method of investigation to help distinguish the differences and, by implication, the relative merit of the methodologies. Harvey's criteria for comparing the four general investigation models were derived from purposes of accident investigations. He judges one model the best. In an earlier ranking of accident investigation models Benner suggested 10 criteria for judging the merits of investigation models of 17 organizations, derived from investigation objectives; statutory mandates; organizations' accident investigation work products; interviews; and previous research discussions. Though their report did not focus on comparing investigation methods, Henderson et al in their study of investigation drivers, methods and outcomes offered criteria for good investigations derived from surveys The Center for Chemical Process Safety has published a comparison of 18 investigation "techniques" with eighteen attributes noted for each. The attributes used are based on "basic concepts, degree of recognition, and areas of application."

This prior work did not compare the methodologies from the perspective of their effects on investigators' specific tasks during investigations. An obvious and probably the most persuasive way to make such a comparison would be to conduct simultaneous investigations of the same accident using two or more different methodologies, and compare the investigation conduct and

outputs. The project would have to be designed carefully to minimize interference between the two investigations. Before expending resources required for such a project, it seemed reasonable to try to explore a less resource-intensive approach for the comparative evaluations, to see if it might disclose differences.

The approach: The approach devised was to use data from an accident report developed with one methodology as inputs to a "desktop simulation" of an investigation with another methodology. Details about the methodology used or how they influenced the investigations tasks that produced a report are not usually reported, except for special tasks like tests or simulations, for example. However, it seems reasonable to assume that what is reported reflects the influence of the methodology used on the investigation and report.

Case Selection: Because of its widespread use in the nuclear, chemical and medical fields, the methodology chosen for the data source is the Root Cause Analysis (RCA) methodology, represented by a report prepared with a variant of that methodology. The methodology chosen for the desktop simulation, the Multilinear Events Sequencing–based (MES) investigation methodology. Both were analyzed in prior studies.

The case selected for this examination is described in a comprehensive accident report by the Chemical Safety and Hazards Investigation Board. [6] It is the policy of the CSB to always do a root cause analysis. An accident in a plastics manufacturing plant fatally injured three workers. The report of the investigation contains a very extensive description of that accident process, the root and contributing causes found, and recommendations to prevent similar occurrences. Although the report does not identify the investigation methodology used, the findings, the time line, the analytical logic trees and language in the report reflect the key elements of RCA, indicating the investigation methodology used was the Agency's variant of RCA.

Description of Methodologies: RCA is an experience-oriented investigation methodology evolved from the U. S. Navy's nuclear submarine program and MORT research performed for the U.S. Atomic Energy Commission. [7] MES is a self-directing rule-oriented investigation methodology, derived from observations of and experimentation with investigation processes at the NTSB and elsewhere.

RCA is a team investigation process. Investigators require extensive training in the method, because its application is heavily experience oriented. After the team is formed, the investigators begin acquiring data, using "why" questions, and use check lists, events and causal factors charts, logic trees, cause trees, and similar aids to identify and document the causal factors. After the causal factors are developed, guides offering categories of basic or root causes are available to help investigators or analysts negotiate, select and report the basic or root and contributing causes, and make recommendations. The narrative report is prepared from these data and findings. Narrative reports are usually accompanied by a chart, a timeline or logic tree to help communicate the investigation results.

RCA has many variants, but the variants have similar goals, and use many similar tools and procedures. The goals are to determine root causes of accidents or occurrences that can be used for achieving future performance improvement. Another common goal is producing recommendations about problems that "management has control to fix, and when fixed will significantly reduce or prevent the problem's recurrence."[8] The RCA investigation tools include events and causal factors charting, time lines, logic tree-based analyses like fault trees, and causal maps or guides to help investigators. Some variants use specialized tools like fishbone diagrams,

Why Staircases, change analysis, barrier analysis, and energy-barrier analysis. RCA software packages are available to implement several of the variants.

MES is a self-directing system of concepts, principles, rules and procedures for any kind of investigation. [9,10] It focuses on identifying, describing, and explaining the behaviors and interactions of people and objects during accidents, and on developing changes in those behaviors that can be implemented and tracked to achieve continuous improvement in future performance. MES views the occurrence being investigated is a process, which can be flow-charted when it is adequately understood. Adapted from the structure and notations for documenting musical scores, MES uses a matrix-based structure with data documentation, organization and analysis tools and rules, plus continuing logic tests, to drive the investigation tasks. Matrix entries follow prescribed grammar and syntax rules of construction for event blocks on matrices including person, number, tense, voice and deictic position -- MES data language -- and reasoning rules to develop tested descriptions and explanations of what happened. MES then analyzes those descriptions systematically with orderly sequential problem defining, change development and ranking tasks and rules. MES provides some generalized behavioral models, guiding principles and assessment or ranking tools to convey knowledge from prior experiences to help investigators. Software implementing MES and self-directing learning tools are available.

The Simulation.

The simulation with the MES-based methodology started with the notification of the accident and creation of an MES matrix, and continued with the acquisition of more data, guided by the data already recorded as event blocks on an expanding matrix. As data were needed, the report was perused to find it and add it to the matrix. If the report did not provide needed data, gaps became visible on the matrix. These gaps were filled with inferred or hypothesized event blocks, and the report was again scanned to find supporting data. If it was not found, the matrix was marked with a "?" or e or dashed arrow or comment to highlight the gap. The evolving matrix guided the investigation effort.

The Simulation Data Source: The 97 page CSB report contains data about the facility: its history and its operation; a diagram of the process equipment involved; a description and analysis of what happened; a description of some aspects of the investigation, a discussion of relevant regulations; the three root causes and four contributing factors found; eleven recommendations; a list of seventeen literature references; a variant of an events and causal factors chart; and several logic trees dealing with specific aspects of the accident. The report does not describe the conduct of the investigation, except in a few instances where it describes tasks related to acquisition of specific data like pressure estimates, so how specific tasks were performed is not known. Thus, while differences will be observed, nothing in this paper should be construed as disparaging the CSB investigation because of those uncertainties.

Assumptions: To keep the focus on the effects of the methodology on the investigation tasks, the simulation assumes that the notification, case selection criteria, capable investigators, accessibility to data and all respective methodological supporting tools are equally available for both methods. It also assumes that all the factual data in the CSB report are true and faithfully reported. Finally, it assumes the first notification was a report within an hour of the accident that three workers had been killed by a vessel explosion in the BP-Amoco facility, and that an investigation team was dispatched from CSB headquarters based on the initial report.

The Simulated Investigation: MES-based investigations use the accident notification and case selection to initiate the investigation tasks.

The company notification informed the CSB that two employees were killed instantly when a tank end cover at its Amodel Plastics Facility blew off, and a third injured employee was pronounced dead on arrival at a local hospital shortly after the accident. The case was selected for investigation.

The first step in launching the simulated MES investigation is to create a matrix or worksheet, on which investigators document and organize data as they become available. In this simulation, information provided in the notification is used to start the documentation, as shown in Figure 1. This document is expanded as the simulation progresses, leading to a 'flow chart" of the entire accident when it is completed.

Figure 1 shows the format of the MES matrix and its contents To permit placement on the matrix, all data used in the matrix must be transformed according to certain grammar and syntax rules into a timed actor + action format, called "Event Blocks" or EBs, the investigator's building blocks. The unique name of the person or object is the actor, required to define the actor row on which to place the EB on the matrix. Actors are the people or objects whose behaviors produced the outcome, and whose behaviors might have to be addressed by any risk reduction proposal. Actions are required to define what the person or object did during the accident process to advance it and to define the temporal or spatial sequencing of those actions along the actor rows. Thus the matrix disciplines investigators and anyone else involved with the investigation to focus on gathering behaviors or actions –Event Blocks. RCA looks for events and conditions.

Figure 1 - Initiation of MES Matrix

The EBs also include other data used by investigators during the investigation. Each contains, in addition to the actor and action, additional attributes of the action, including the location where the action occurred, the source(s) of data from which the action was defined, the time the action

began, and the time it ended. Investigators may also add reminders or notes about possible further investigation tasks or encoded colors for highlighting different kinds of Event Blocks like incomplete EBs, disputed EBs, inferred EBs, etc. RCA has not comparable requirements.

The grammar and syntax rules help overcome five common investigation problems observable during investigations and in reports:

1. Using more than one name for a person or object in the report which confuses users as they try to follow the interactions during the accident. (CSB report. page 23 uses the terms "prepolymer" and "reactor effluent" for apparently the same material.)
2. Using plural names like "the crew" or they or pronouns, which prevents the mental visualization of the accident, raising questions about the investigation. (CSB report uses many plural ambiguous terms like operators, workers, Supervision, etc.)
3. Masking incomplete or inadequate investigations by reporting actions in the passive voice, which obscures what happened, and the specific actors and behaviors that need to be changed to reduce future risk. (CSB report page 27 contains a dozen examples.)
4. Introducing assumptions, opinions or generalities unsupported by evidence acquired during the investigation into the accident description. (CSB report had none.)
5. Overlooking or not reporting behaviors that influenced the course of the accident and outcomes. (CSB report omissions are described below.)

MES requires a source for every EB. This means investigators must develop a description of what happened from observed accident data, and report when experience or other sources are used to create EBs. This source requirement also provides a tool for inventorying, retrieving and managing source materials accumulated during the investigation. RCA does not require and the CSB report does not indicate the source of all events described, so there is no way to confirm their validity, particularly in the sections describing operations just before the explosion.

The EB "begin time" requires investigators to document the timing of behaviors during the accident process when it is know, or to estimating those times to assist in the ordering of the EBs in the matrix. RCA leaves time documentation to the investigator's discretion.

The EB placement procedure on the matrix provides investigators a form of "progressive analysis" by applying four kinds of logical reasoning to the data as are added: a) sequential reasoning, for ordering EBs, b) cause-effect reasoning for establishing causal coupling, c) deductive reasoning for deriving or inferring EBs from data or bridging gaps in the event flows, and d) necessary and sufficient reasoning for determining the completeness of a description and explanation of the accident. The arrows on the matrix are a way to quickly show tentative couplings among EBs as properly sequenced new EBs are added to the matrix. RCA has no formalized linking or testing rules.

The RCA events and causal factors charting, one of the principal RCA tools, offers investigators a somewhat similar though substantially less rigorous progressive analysis capability. It uses conditions as well as events selected at the discretion of investigators, accepts ambiguous entries, often lacks timing relationships and grammar or syntax rules, and requires fewer logical validation tests, thus requiring more subjective decisions of investigators.

Building the matrix: Continuing the simulation, the MES matrix helps define what to look for next. On arrival at the scene, investigators typically view the scene, and try to find witnesses who can report their observations about the occurrence. In this case, the scene offered additional EBs. These EB were taken from the report, which is assumed to have resulted from what investigators

reported about the objects involved and the damage to the surroundings. For example by viewing the debris, investigators would observe that the polymer catch tank (PCT) end cover had blown off the tank; 22 (50%) of the bolts holding the cover in place, from about 7:30-1:30 on the flange, had broken; the cover struck a canopy sheltering the catch tank and landed 14 feet from the PCT; the contents of the PCT including the metal frame were expelled from the tank; and apparently, based on the position of the victims after the accident, struck the victims. It was not clear from the report whether the cover had also struck any of the victims. (The report contained an internal ambiguity on this point on pages 10 and 30.)

Adding these data to the initial worksheet, using the information from the report, one can see several more effects of the MES methodology. See Figure 2.

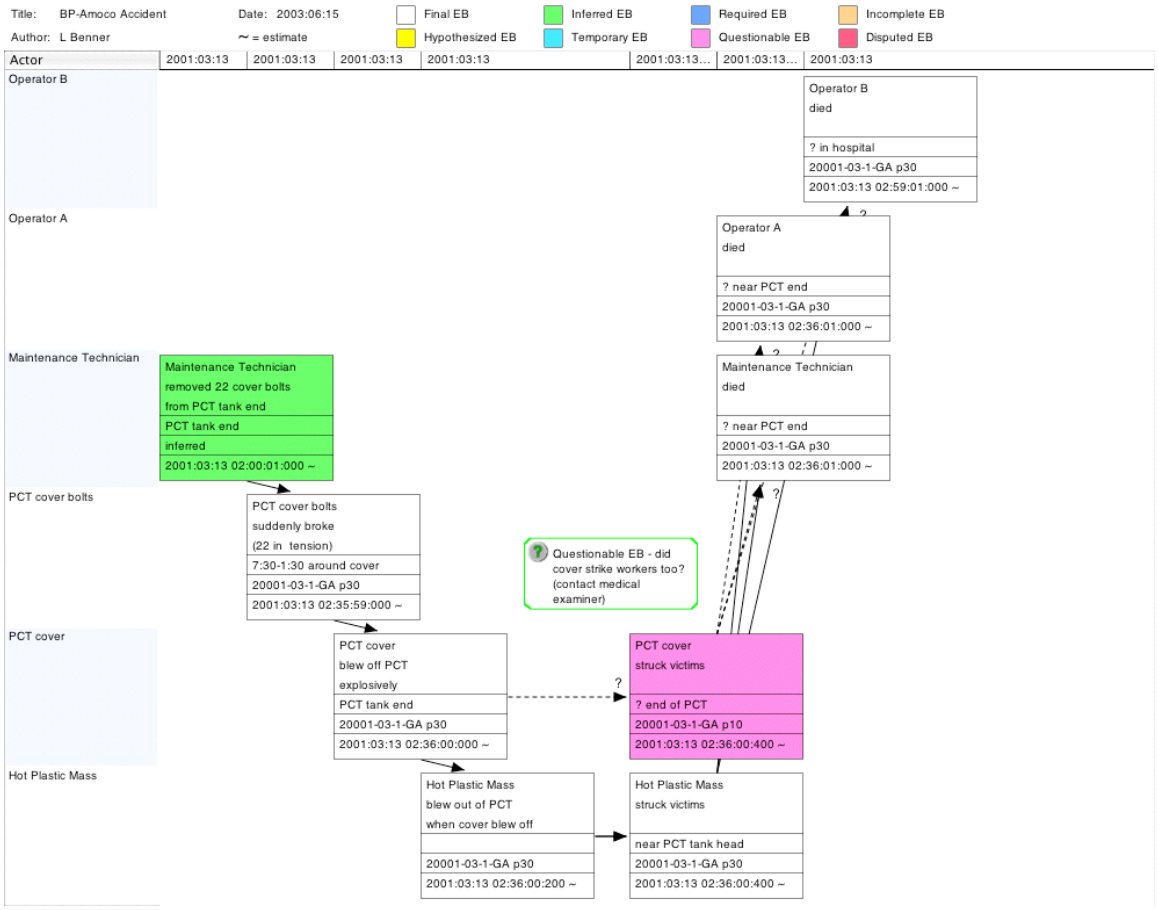


Figure 2 - Adding Data To An MES Matrix

This shows MES features affecting what investigators do and how they do it during investigations. On arrival at the site where the victims were killed, and after talking to the supervisor, investigators found additional EBs, and added them to the matrix. Two of the new EBs are color coded: the first (green) indicates to all working on the investigation that the EB is logically inferred by the investigators, and the second (cyan) indicates that that there is some question about the EB which should be resolved before the investigation is closed. The arrows were added between EBs that have a logical cause-effect or “this had to happen for that to happen” relationship during this accident.

The inferred EB illustrates how EBs evolve during the investigation. The investigator's observation that 22 bolts holding the end cover had broken indicates that the maintenance technician had removed the other 22 bolts before the cover blew off, and may have been working on the 23rd bolt when the cover blew off. The broken bolts, together with the distance the cover flew and its trajectory after it broke away from the PCT, enable investigators to infer that a substantial internal pressure inside the PCT drove the cover of the PCT to where it went. While this seems intuitive, the EBs provide *anchors* for this reasoning process. Thus the matrix helps investigators pinpoint what to look for next, and every investigator involved in the investigation can access that knowledge at any time. Good investigators do this intuitively but informally; the matrix helps investigators capture, document and use their observations.

The report is silent on where the victims were found and the results of autopsies if they were performed. The report says they were struck by the hot plastic mass that blew out of the PCT after the cover blew off. Therefore investigators show this cause-effect relationship with a solid arrow. Because of an inconsistency in the report, there is a question about whether the end cover struck any or all of the victims, so the investigator shows a dashed arrow to indicate another open investigation data item to resolve. Dashed lines, question marks and comments show everyone all the uncertainties or problems with the accident description.

Investigators add links between EBs that are logically coupled. The links result from investigators' spatial and temporal sequential reasoning and cause-effect reasoning. MES also requires necessary and sufficient (N/S) reasoning as the EBs are added, to determine the completeness of the links to an EB. For example, applying N/S reasoning to one of the EBs, "Hot Plastic Mass struck victims," we find that more EBs are needed before that EB can be replicated.

MES rules require investigators to reason their way through the N/S completeness testing procedure, to help them identify all the predecessor behaviors required to produce the EB. The decision steps are supported by specific questions the investigator must answer. [11] When the EBs on a Matrix form a linked path from beginning to end, the description and explanation are complete. If gaps remain, they must be explained. Unlinked EBs are analyzed to determine if they are relevant; if the are not, they may be removed.

For the hot plastic mass to strike the victims, the victims had to position themselves in a place where the mass would strike them when it blew out of the PCT, and it had to happen so quickly that time for them to escape its path was not available. This leads investigators to pursue the reasons why the victims did that, which leads to further questions about procedures, tools and how the tools and procedures came into being. Answers are added as new EBs. These questions are continued until the investigator is satisfied that all the preceding EBs needed to reliably reproduce the trailing event each time they occur are identified and added to the matrix. For example in this case, the victims also had to remove some bolts in a certain way for the cover to blow off and let the plastic mass blow out of the PCT. The observed dispersion of the hot polymer mass indicates that the mass was expelled from the tank under considerable pressure, even though the tank was thought to be empty. Upon realizing this, investigators would add a temporary EB about something pressurizing the interior of the PCT. Pursuing this kind of reasoning can lead investigators to procedures, training, design, policies and other actions that were necessary for an EB to occur. These events would be added to the matrix, as shown in Figure 3.

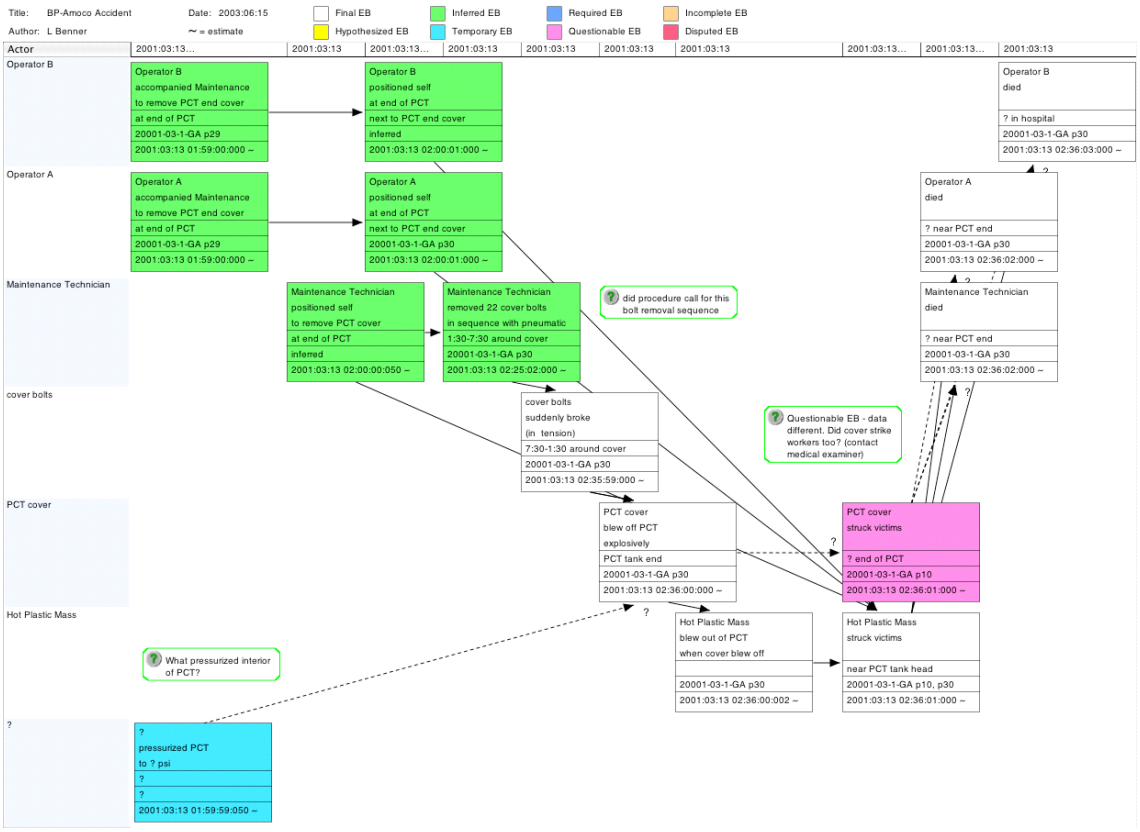


Figure 3 - MES Matrix Showing Additional Actions

The CSB report's "Timeline" describes the injury events with the first three blocks in Figure 4.

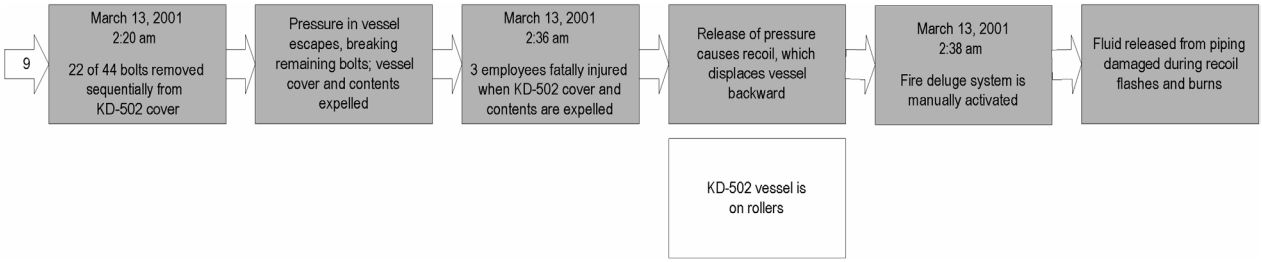


Figure 4 - CSB Timeline Of Worker Injury (page 92

RCA logic trees reflect similar kinds of reasoning about events, but the trees lack the grammar, syntax and N/S logic testing and the disciplined thinking they demand. In the absence of RCA rules for when to end logic trees, inquiry into procedures and their development depends on the investigator's judgment.

Interestingly, although the CSB report describes another accident it had investigated where 6 victims were killed when a tank cover was opened, the published graphics do not show these points. This affects the recommendations. Investigators using MES would explicitly describe both aspects in the matrix. The differences are attributable to the tasks required of investigators by the methodologies. The consequences can be observed in the differences in the outputs.

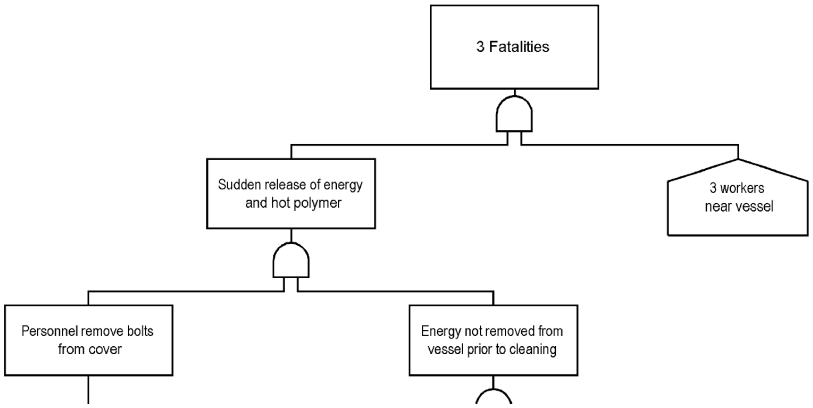


Figure 5 - Logic Tree of Worker Injury (page 93

Investigation depth: In this simulation, the part of the MES matrix developed from HAZOP data in the report demonstrates more differences. Figures 6A.1 and 6A.2 show the CSB results, and Figure 6B shows how the MES would require further investigation effort. The CSB report addresses the hazard analyses that were performed to develop safety controls for the process. The report's time line shows the problem as Less Than Adequate PHA (Process Hazard Analysis).

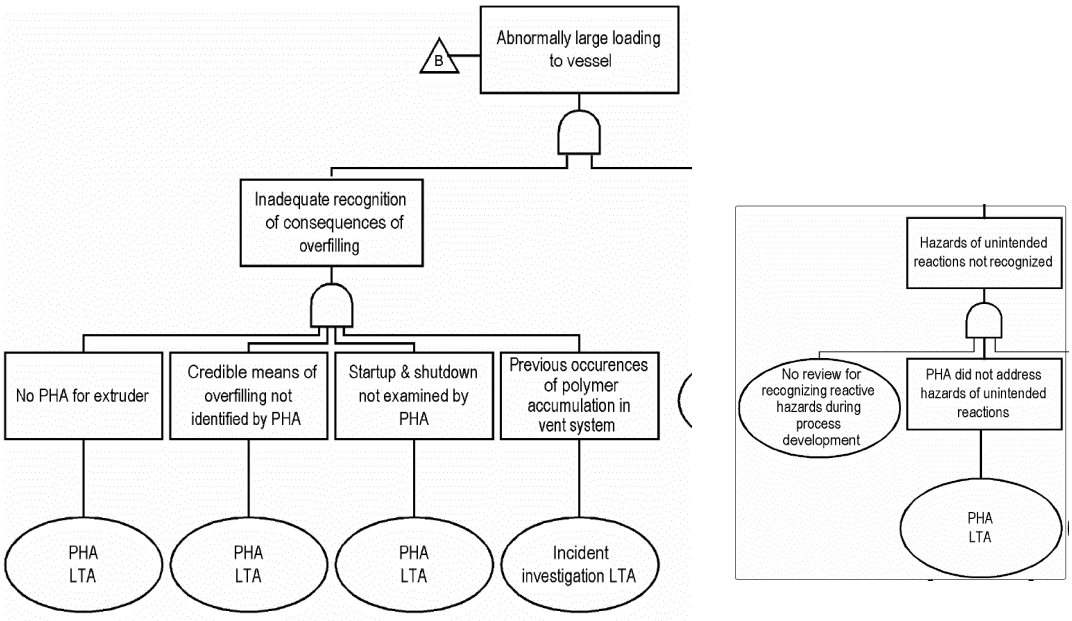


Figure 6A.1 - CSB Logic Diagrams of PHA Roles, (page 95-96)

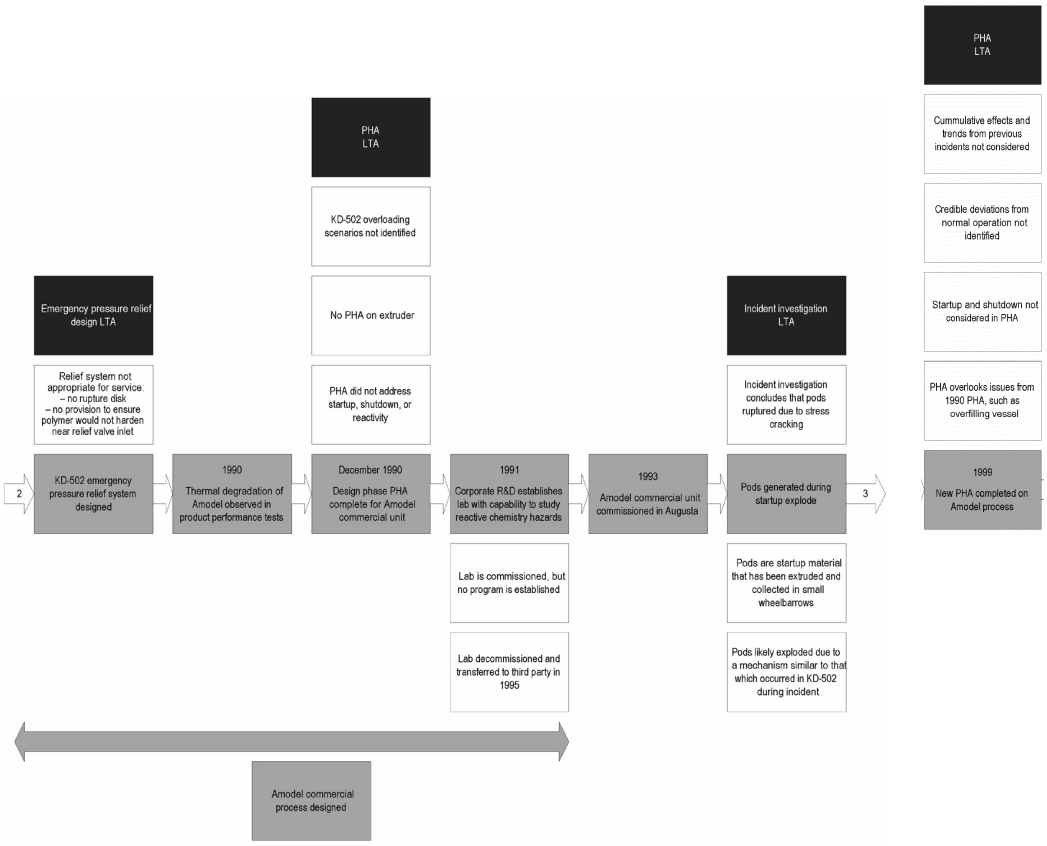


Figure 6A.2 - CSB Timeline of HAZOP Role (page 86)

The report also contains a logic diagram for the abnormally large loading of the vessel, showing the Preliminary Hazard Analysis (PHA) as an unanalyzed “basic event.”

Upon discovering that a HAZOP analysis was used for analyses of process hazards, MES investigators would try to ascertain whether the analysis method or its implementation resulted in the overlooking of the hazards or unsuccessful hazard control in this operation. HAZOP is a widely used analysis method in the industry and the difference would be important. Investigators using MES would display what happened somewhat differently. Figure 6B, abbreviated for this paper, shows the role of HAZOP analyses of the process, and shows how the EB array leads to several investigation tasks to clarify that role. Testing the EBs with necessary and sufficient reasoning leads investigators to the assumed, temporary and inferred EBs shown here, and some of them would, in turn, lead investigators to still others not yet shown. The reasoning leading to further investigation tasks is documented by the temporarily assumed EBs (blue) and the inferred EBs (green) and the questionable EB (cyan). Each indicates unfinished investigation tasks. This process could lead to management, technical or regulatory links, depending on what investigators find as they work their way upstream of the EBs shown in Figure 6B.

Incidents between the 1990 and 1999 HAZOP analyses which were not reflected by the 1999 analysis raise further questions about the analysis method and its implementation in the safety and management system; this would require further investigation to ascertain what happened during the analyses and tracking functions, and why it happened.

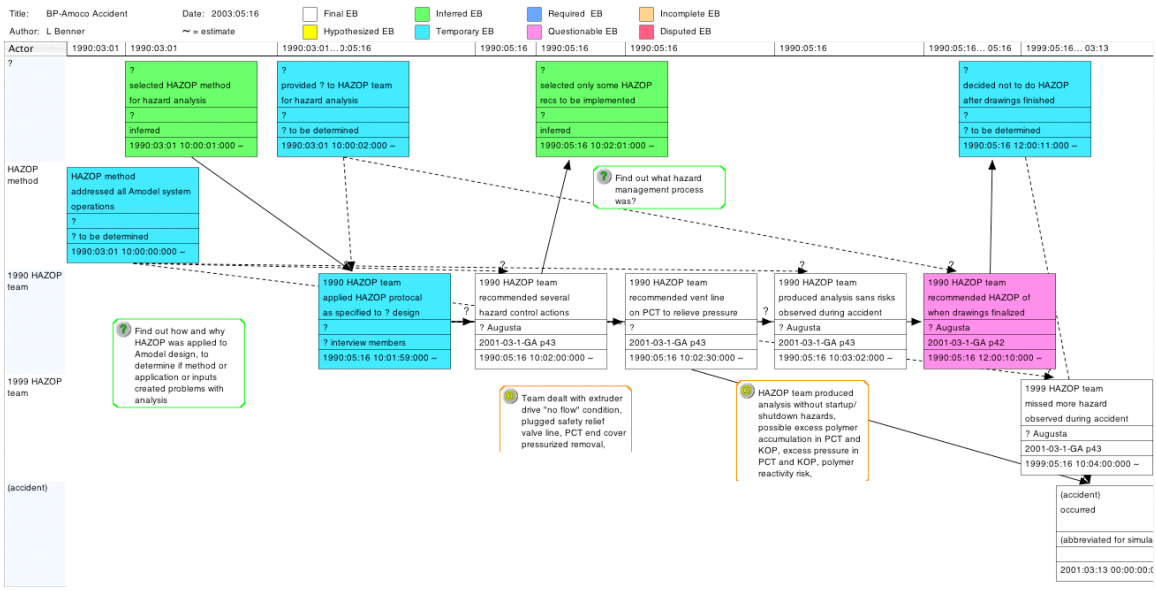


Figure 6B - MES Matrix re HAZOP Role

This significant aspect of this accident was addressed generically by the CSB report, to its credit, but it the report is unclear why the problem occurred and whether the recommendation that followed is properly directed or will achieve needed behavioral changes to reduce future risks. Further, since the desired actions are ambiguously defined, the ability to verify their successful implementation will depend on the occurrence of future incidents, rather than observations of specified behaviors during ongoing operations.

The simulated investigation continues in this manner until the origins of the accident are identified, documented and coupled, to provide as complete a description and explanation of the accident as is possible from report. As the MES investigation proceeded, it raised numerous questions, which are recorded as comments on the Matrix for disposition.

Recommendation development task: The RCA approach to defining problems is to define root causes and causal factors, which are any problem with the incident that if corrected could have prevent the incident from occurring. In RCA, the root causes are considered to be the problems demonstrated by the incident, and recommendations for remedial actions address those causes or cause factors, as was done in the CSB report. RCA provides various kinds of guidance for development of recommendations, ranging from check lists and templates to computer software.

In the BP-Amoco accident, the CSB selected three root causes: the developer's inadequate review of the conceptual process design to identify chemical hazards, the facility's lack of an adequate review process for correcting design deficiencies, and the site's inadequate system for investigating incidents' and near miss incidents' causes or related hazards. It selected four contributing causes: inadequate and incomplete hazard analyses of the process, inadequate description of the process in design documents, equipment operating procedures that did not specify what actions to take when safety precautions could not be met, and not subjecting revisions to operating procedures to management of change reviews to evaluate safety effectiveness. These causes were abstracted from the description of what happened, starting with the original development of the process until the last fire was extinguished.

The MES methodology does not use causes. Instead, it focuses on specific behaviors and interactions that need to be changed to improve future production process performance. MES strategy for identifying problems or needs is to try to identify interactions that should not occur or should occur differently to change the process flow. Investigators examine each of the behaviors or interactions which advanced the accident process and for each, define any problem it suggests. Each EB and each linked pair or set of EBs is reviewed as a potential problem, and a potential candidate for intervention in the accident process or production process. Each EB pair offers a candidate behavior to look for in future activities, which can then be monitored to ascertain whether the objectionable behaviors or interactions recur after changes are implemented. The HAZOP examples show how that would lead to different recommendations.

Discussion of results

Examples of differences observed already demonstrate that methodologies affect investigations. Are these differences important? Clearly they are, because they lead to different insights, work products, safety issues and recommendations. Their full extent is being determined. When the project is completed, a summary of the all the differences found will be prepared.

References

1. Sklet, Snorre, *Comparison Of Some Selected Methods For Accident Investigation*, NTNU/SINTEF, Trondheim, Norway, 2003

2. Harvey, Michael D., *Models for Accident Investigation,* Alberta Worker Health, Safety and Compensation, Alberta, Saskatchewan, 1985

3. Benner, L., *Rating Accident Models And Investigation Methodologies*, Journal of Safety Research, 16:3, Fall 1985 Chicago, IL

4. Henderson, J., Whittington, C and Wright, K., Accident investigation - The drivers, methods and outcomes, HSE Report 344/2001, Human Reliability Associates, 2001

5. CCPS Center for Chemical Process Safety, *Guidelines for Investigating Chemical Process Incidents*, American Institute of Chemical Engineers, New York, NY 1992

6. U.S. Chemical Safety and Hazard Investigation Board, Investigation Report: *Thermal Decomposition Incident (3 Killed). Report No. 2001-03-1-GA, June 2002.*

7. Johnson, W. G., *MORT Oversight and Risk Tree* , SAN 821-2, U. S. Atomic Energy Commission, 1973

8. Paradis, M. and Unger, L., *TapRoot*^{R}, System Improvements Inc., Knoxville, TN 2000

9. Hedrick, K., and Benner, L., *Investigating Accidents With STEP*, Marcel Dekker, 1986

10. Benner, L., *10 MES Investigation Guide*s, Starline Software Ltd., 2002

11. Benner, L.,*10 MES Investigation Guides, Guide 2 Task Guidance For Organizing And Analyzing Investigation Data*, Starline Software Ltd. 2002

Biography

Ludwig Benner Jr Starline Software Ltd. Oakton, Virginia U.S.A., telephone 1.703.620.2270, e-mail: luben@starlinesw.com

Benner is a retired NTSB investigator, enjoys Scrabble, being www.iprr.org webmaster, and keeping up with investigation process research.

Evaluation of Navigators' Performance Shaping Factors in Marine Incidents

Yoshio Murayama; Maritime Labour Research Institute; Tokyo, Japan
E-mail: JDV00353@nifty.ne.jp

Yusuke Yamazaki; Toyama National College of Maritime Technology; Toyama, Japan
E-mail: yamazaki@toyama-cmt.ac.jp

Keywords: Incident, Marine, Performance shaping factor

Abstract

This article discusses progress in the development of an incident investigation system to improve maritime safety management. The authors designed a questionnaire about marine incidents and the conditions of navigators' Performance Shaping Factors (PSF) in those incidents. The results revealed important dangerous relationships between unsafe acts and PSF, through analysis by a contingency table and stratification of the contingency table by a third factor, step by step according to cause and effect. The results showed that an unsafe delay in the recognition of danger was strongly affected by 5 factors directly, 13 factors indirectly, and some background factors. Mistakes in decision-making were affected by 5 direct factors, 19 indirect factors, and a few background factors. The frequency of these unsafe acts ranged from 11 to 40%, depending on the combinations of the factors. The particularly influential combinations of factors contributing to delays in danger recognition were 'Enthrallment' with 5 indirect factors and 'Drowsiness' with 5 indirect factors. Those of mistakes in decision-making were 'Drowsiness' with 10 in direct factors and 'Unexpected' with 5 other factors.

Introduction

The oil tanker *Exxon Valdez* grounded on Bligh Reef off Alaska in 1989 and spilled a large amount of crude oil, polluting the pristine Prince William Sound. It was one of the most highly publicized accidents in the history of marine pollution. This oil spill highlighted various safety issues in maritime traffic and led to studies by the International Maritime Organization (IMO), which governs international maritime traffic. Until now, the IMO has adopted a number of specific safety measures but has failed to come to a conclusion concerning matters of seafarers' fatigue. Its failure to reach a consensus on this issue is mainly attributable to a lack of a clear understanding of the relationship between the degree of fatigue and the occurrence of marine accidents, or of what contributes to the fatigue of a seafarer and to what degree.

In order to clarify such matters, the IMO adopted resolutions to standardize both a method to investigate marine accidents and international cooperation in such investigations, with special emphasis on human factors. These resolutions have expanded the scope of investigation to include hazardous events that might have led to casualties (IMO,1997), and they require investigation on safety management (IMO,2000). The investigation needs various point of view; Reason's defense model (Reason 1994) and Hawkins' SHEL model (Hawkins 1992). Every country is currently studying specific ways to investigate marine incidents. In the private sector, various incident investigations have already been conducted. It is difficult to use their results for public purposes, however, since they are not standardized. It is therefore necessary to develop a practicable, standardized method to investigate marine incidents and a method to analyze such data for safety measures.

Against such a background, we have developed a method to investigate human factors in marine incidents and to analyze the collected data. In order to clarify the factors that contributed to actions (performance shaping factors or PSF; Miller, 1987), we have devised an investigative method for navigators who have experienced marine incidents. This method includes a checklist to investigate not only events but also the PSF involved, i.e., the conditions of the navigators, facilities, environment, and management (Murayama, 1999).

The analysis of collected data concerning collision incidents clarified problems in the recognition of a dangerous positioning of a ship relative to another ship, particularly with regard to the potential for collision. In addition, a contingency table, an extremely fundamental methodology, revealed important relationships between the events and their causal factors, which we call direct factors (Murayama, 2002).

This article reports on a method to identify the indirect factors that influence the direct factors in maritime incidents, as well as background factors that influence the indirect factors. The article provides a method to evaluate the influences that cause the events.

Concept of the methodology

Links between events and factors: Safety measures to prevent accidents are derived from investigations into the causes and contributing factors of accidents that have already occurred. For such investigations, fault-tree analysis (FTA) and event-tree analysis (ETA) are used, following the time sequence or causal sequence of events and factors. FTA and ETA are rather easy to structure when the order of operations, or the difference between normal and abnormal states, is clear, as in the case of industrial plants. But both types of trees become complex and vague when operations greatly differ and/or multiplex selections are possible, depending on external conditions.

Human action, in particular, is affected by numerous factors and the relationships among them. For example, when two ships encounter each other in marine traffic, they can select between two maneuvering actions: (1) if there is ample distance, the ship can dissolve the positional relationship before marine traffic rules bind it to a specific course; or (2) when the distance is not ample enough, the ship can avoid collision by following marine traffic rules. This selection is linked to several PSF: the sea area, traffic, the navigator's cautiousness, time pressure, etc. It is difficult to construct a tree that links these factors.

However, bad PSF conditions do not always result in unsafe acts or incidents, because ordinary working conditions vary. Likewise, although there are various unsafe actions and conditions in PSF, a factor is not necessarily related to an unsafe action, or an unsafe action to an incident. Consequently, we need to evaluate the influence of PSF on unsafe actions and incidents by comparing the frequencies of unsafe actions in every type of incident and under each PSF condition.

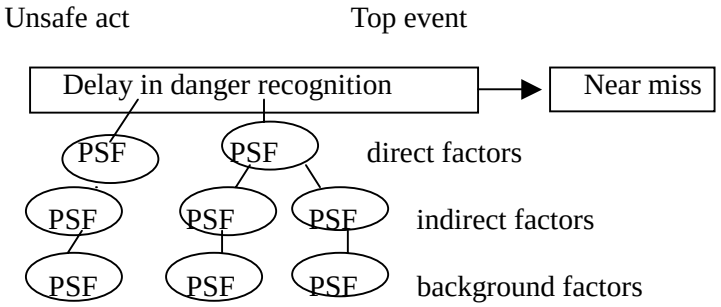


Figure 1 - Links between events and PSF

Relating factors to problems: To estimate the influences that PSF and PSF conditions have on human actions, psychologists and sociologists often rely on a coefficient of correlation matrix or a score of a principal component analysis.

These indexes mainly emphasize deviations and the numbers of subjects. They disregard small numbers of cases as special unsafe acts and discount PSF as important components of an accident. Furthermore, the value of a principal component analysis changes every time a different combination of PSF is studied.

The simplest approach is the use of a contingency table, which reveals…categories, combined with a specific method to show a one-to-one correlation between an event and a factor or between factors. The odds ratio of a contingency table based on the number of cases is not much affected by the total number of cases. For these reasons, it is easy for a businessperson to use the table and to understand the results. This approach is effective for revealing relations when the number of cases is small. In addition, a contingency table created by combining other factors allows us to examine correlations with multiple factors.

Procedure of the method

Target of analysis: The first step in the analysis is to categorize the problem into an incident type, based on the frequency distribution of similar incidents. We also clarify the PSF involved in the incident; we identify these direct factors by a contingency table between the problem and the PSF.

Factors involved in the target: Next we clarify what we call indirect factors, which are the PSF that contribute to direct factors. For this, we use a contingency table between the direct factors and the PSF. Then, in turn, we clarify the background factors, which are PSF contributing to the indirect factors. For this we use a contingency table between the indirect factors and the PSF. These relations are selected based on an odds ratio of over 1.3 and are confirmed by Yule's coefficient of association.

Effects of third factors: Some relations may be strongly affected by other factors (third factors). For this reason, when we evaluate one-to-one relations between a problem and the factors, it is also necessary to consider the impact of third factors on those relations. The contribution of third factors to a problem can be evaluated by stratifying a contingency table by the third factor between the problem as a dependent value and the factor as an independent value. When we consider the occurrence of a problem, two contingency tables between dependent values and independent values are described as the two lines in the Figure 2.

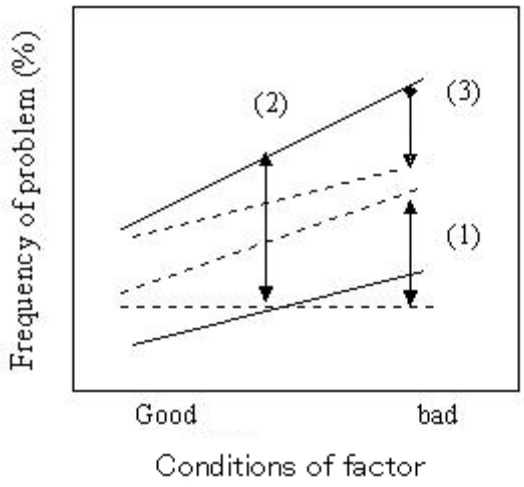


Figure 2 - Evaluation indexes of third-factor effects

When the effect of the independent value is large, the inclination of the two lines is large. In order to neglect the influence of third factors in the relation, we consider that the inclination of the middle line between the two lines is the basic component representing the direct effect ((1) in the figure).

When the two lines diverge to the upper and lower sides, it means that a third factor has a large influence on both lines. We consider that the gap in the frequency of the dependent value between the of the lines is a weighting component representing a separation effect ((2) in the figure).

When the line segments cross each other, it means there that a third factor is causing a large difference between the groups. We consider that the difference between the inclinations of the lines is a cross component representing the interaction effect ((3) in the figure).

We disregard any relations that have differences of less than 5 percent in the frequencies of problems of dependent values between the divided groups.

The relations between incident events and PSF conditions become complex and redundant according to progress analyses into the factors.

Order of cause and effect: We coded factors into two types in order to reduce the need to analyze relations. One type of code is that of the causes and effects of the events and of the PSF. These include: (1) matters concerning an unsafe act that occurred during an incident (2) matters directly related to those acts (direct factors), (3) matters related to those direct factors (indirect factors), and (4) matters in the background (background factors). This code is used to create a set of factor chains showing causal relations, as in a fault tree.

The other types of codes are of the fields of safety measures to concentrate on the practical use of the results. This code are for PSF and for SHEL that are categories of safety management.

Collecting data

The data of this report were in the form of answers to our questionnaire about situations surrounding marine incidents and the PSF conditions of navigators who have experienced a dangerous situation. The questions were 100 items: 55 main items included an additional 45 sub-items for PSF in seven sections. The answers were in the forms of numerical values and adjectival scales. For example, the value of the recognition time of danger was "the time from the moment when danger was first recognized until the most dangerous distance to the obstacle was reached"; relations to an obstacle ship were "crossing from the right", "crossing from the left", "overtaking", or "facing"; and the fatigue scale ranged from "normal" to "fairly poor", "poor", and "very poor".

The subjects of these investigations were the bridge navigators of 2,351 coastal ships belonging to 167 companies involved in domestic maritime transport. The questionnaires were mailed to the captains and distributed to the subjects through the safety managers. After filling out and sealing the questionnaires anonymously, the subjects submitted them to the captains, who forwarded them to the respective safety managers. The sealed questionnaires were then returned to the researchers. We got 2615 responses about incidents from 2831 respondents on 1274 ships (45% of subject ships).

Results

Two specific problems emerged from the frequency analysis. One was that by the time two ships become aware of each other, the recognition of the danger of collision comes too late, in many cases, for the appropriate action to be carried out. The other was that the ship's decision-making in relation to the other ship's movement was unsuitable.

It is possible to grasp the delay in recognizing a danger from answers to questions contained in the checklist, specifically those concerning the time that elapses from the recognition of a danger to the occurrence of a most dangerous situation (danger recognition time). Close to half of the respondents answered 'less than 1 minute'. This is not enough time for a general merchant ship engaged in domestic trade to change her course by 90 degrees. This is regarded as a delay in recognizing a danger.

Mistakes in decision-making can be identified from answers to questions concerning the position of a navigator's own ship in relation to that of another ship. A dangerous situation, in which another ship crosses the navigator's course from the right, is a situation in which the navigator, in guiding a give-way ship, has failed, and it is regarded as a mistake in decision-making on that navigator's part. Conversely, if a dangerous situation has occurred and the navigator's ship is a stand-on ship, the other ship has made a decision-making mistake. The frequency with which a navigator's own ship acted as a give-way ship was 30 percent.

In order to study measures to prevent collision by these problems, we clarify whether or not factors contributed to or facilitated the occurrence of such unsafe acts, and if so to what degree.

Delay in recognizing danger: Problems in recognizing danger; either under 1 minute or over 1 minute from initial recognition until the most dangerous moment passes, was related to 21 direct factors**.** These 21 direct factors were related to 353 indirect factors; the number of the factors was redundant total of them. These 353 indirect factors were related to 4991 background factors. Excluding the relations of the factors that were inversions of cause, the total number of combinations was 1008 cases. We evaluate effect of indirect factors; a separation effect and a

cross component between groups of indirect factors, and selected 170 combinations for PSF of live-ware as SHEL model. The difference in frequency described in this table is the difference in the occurrence of the incident by delay in recognizing danger, the difference between maximum frequency and minimum frequency of the unsafe act that have occurred, and combining direct factors and indirect factors.

Table 1 lists 5 direct factors for PSF in the combinations relating to other factors. The enthrallment relates to 5 indirect factors and many background factors. The other differences of the combinations are from 11 percent to 40 percent. relations between differences of the relations between the combinations were from 16 percent to 33 percent. The background factors expand the differences.

Table 1- Difference of frequency of unsafe act as delay in danger recognition

(unit: %)

Direct factor	Indirect factor	Frequency of delay of danger recognition		
		Min.	Max.	Difference
Enthrallment	Overcrowded schedule	44	64	20
	After strain	45	72	27
	Straying mind	43	69	26
	Anxiety for private life	39	79	40
	Anxiety for work	44	70	26
Unexpected	Anxiety for private life	39	57	18
Fatigue	Bad physical condition	39	50	11
	Anxiety for private life	36	56	20
Drowsiness	Overcrowded schedule	41	57	15
	Bad physical condition	38	56	18
	Straying mind	39	63	24
	Age	32	58	26
After strain	Ability of work	43	61	18

Decision-making: Seven factors were related to problems in decision making while under duty to avoid another ship or to maintain the course of one's own ship. This unsafe act related to 30 direct factors, and the factors were related to 642 indirect factors; total number, and related to 8816 background factors. There were 1055 combinations of the factors selected according to order as cause and event. Combinations of factors of PSF for live-ware were 74 cases. Selected combinations by considering causes of factors for unsafe act are in the table 2. These combinations related to many background factors, which expand the differences.

Value of the tables: Much difference of frequency of every combination in above the tables means indexes of efficiency of safety measures. If we change worse conditions of the factors of the combinations, we can expect to reduce occurrence of unsafe acts by the difference. However there are many complicated factors, and false relations as cause and effect, we have to examine relations that are revealed by this analyses.

Table 2 - Difference of frequency of the unsafe act as mistake of decision-making.

(unit: %)

Direct factor	Indirect factor	Frequency of miss of decision-making		
		Min.	Max.	Difference
Unexpected	Maneuvering style	37	53	16
	Attitude to obstacle	33	58	25
	Many fishing boats	32	65	32
	Rare meeting ships	40	73	33
	Age	37	65	29
Drowsiness	Age	37	64	27
	Job ranking	37	65	28
	Communication of crew	39	62	23
	Ability of crew	40	64	24
	Job ranking	30	62	32
	Type of obstacle	35	53	18
	Methodical	38	58	20
	Carefulness	41	67	26
	Team work	40	62	21
	Maneuvering style	38	63	25
Straying mind	Type of ship	30	61	31
	Size of ship	36	75	39
Anxiety for work	Number of deck crew	36	52	16
Time pressure	Type of labor contract	39	54	15

Conclusions

The method for analyzing navigators' performance shaping factors in marine incidents of collision and grounding is to first identify correlated factors according to items on the incident questionnaire, by employing a contingency table and calculating the odds ratios. The second step of the analysis is to delete certain combinations of factors based on the order of the factors in cause and effect. The third step is to neglect the relations affected by third factors or those affected by third factors that do no affect the problem. Finally, we use the obtained combinations of factors to compare the frequencies of incidents involving combinations of factors selected in the field of safety measures.

In collision incidents, unsafe act as delay in danger recognition was related to 5 factors; navigators' enthrallment, unexpected for ship motion, fatigue, drowsiness and past strain. These direct relations are indirectly affected by 13 other PSF and many background factors. Maximum difference of the frequencies of the unsafe act was 40 percent. In the case of a mistake in decision-making, this unsafe act was directly related to PSF; unexpected for other ship's action, Drowsiness, Straying mind, Anxiety for work and time pressure. These relations affected by other PSF, and maximum difference of frequency of the unsafe act was 39 percent.

This method reveals a tree that simply arranges combinations of many dangerous events and their

influencing factors, and the contingency tables show differences in the frequency of problems among combinations of related factors. A businessperson would be motivated by this analysis, since it reveals the apparent structure of unsafe factors and the effects of those factors on dangerous events.

Reference

Hawkins, H. F. (Kuroda, I. Jpanese ed. 1992) 1992 Human Factor in Flight. 7-11, Seizan-do, (Tokyo)

IMO.1997. Resolution A.849(20). IMO, (London)

IMO. 2000. Resolution A.884(21). IMO, (London)

Miller, D. P. and Swain A. D. (Solvandy, G. ed.). 1987. Handbook of Human Factors. Wiley InterScience. Prudue Univ.

Murayama, Y. et al. 2000. The Journal of the Japan Institute of Navigation, Nautical Institute, 102, 173-181.

Murayama, Y. and Yamazaki, Y. 2002a. The Journal of the Japan Institute of Navigation, Nautical Institute, 106, 55-62.

Murayama, Y. and Yamazaki, Y., (Johnson C.W. ed.). 2002b. Workshop on IRIA2002, 102-114, Glasgow Univ. (Glasgow)

Reason, J. (Shiomi, H. Japanese ed.. 1999) 1994. Managing the Risk of Organizational Accidents., 14-15. Nikka-giren. (Tokyo)

Yamazaki, Y. and Murayama, Y. 2001. The Journal of the Japan Institute of Navigation, 104, 173-178.

Divergence and Convergence, Trends in Accident Investigations

John A. Stoop; Delft University of Technology, Faculty of Technology, Policy and Management;
Jaffalaan 5, P.O. Box 5015, Delft, 2628 BX Delft, the Netherlands
johns@tbm.tudelft.nl, stoop@kindunos.nl

Keywords: accident investigation, methodology, safety boards

Abstract

Transport Safety Boards have seen an incremental development over the past decades, evolving from disciplinary courts and governmental investigation committees into independent safety agencies. New missions, roles and responsibilities have caused a shift in focus in accident investigation, which may lead to a divergence in forms and procedures across such investigations. Simultaneously, a convergence takes place with respect to the establishment of independent safety boards, necessitating the development of a specific methodology. In this methodology, a combination of initiating investigations, fact-finding, safety analysis, drafting recommendations and initiating systems change takes place, defining these agencies as problem providers for safety management and engineering designers.

This contribution discusses methodological issues involved in this evolution in accident investigation from a perspective of Transportation Safety Boards.

Introduction

Maritime accident investigation courts were established by the second half of the 19th century in most of the sea-going trade nations. A judicial approach enabled disciplinary action against the misconduct of a captain and officers endangering vessels, cargo and passengers. The role of the government was exclusive: the findings of the boards were addressed to the ministry, which held jurisdiction over the issue. In most cases this was the ministry of transportation. The inspectorates of the ministries, which also issued the reports on which boards could base their decisions, conducted the investigative efforts. Similar administrative investigation agencies were established in the railways in may countries, although the disciplinary aspect was less prominent or even abandoned for the benefit of learning. Developments in aviation were slightly different from the maritime and railway sector. Accident investigation into major air crashes was established mandatory as an international obligation of a state by ICAO under Annex 13 in 1951. Initially, the focus was on the technical reliability of the aircraft, the performance of the pilot and compliance with regulations.

In the sixties of the previous century, the concept of independent and permanent investigation boards was adopted in other modes of transportation as well, leading to establishing multi-modal transportation safety boards throughout the world [1], [2], [3].

Over the past decade, several major events have occurred across Europe dealing with infrastructure related disasters. Public and political concern has been raised about fires in the Channel Tunnel and tunnels in the Alps region, high speed train crash at Eschede in Germany and a series of railway accidents in the UK, capsizing of the passenger ferries Herald of Free Enterprise and the Estonia, grounding and sinking of sea-going crude oil tankers, the Concorde crash and the mid-air collision over the border of Germany and Switzerland. In the aftermath of these events, questions have been raised about the preparedness for such disasters and capacity for emergency response, salvage and rescue. Consequently, a need for prevention, policy

harmonisation and regulation at a European level has been identified. Draft Directives are prepared in the European Union to establish mandatory safety agencies and modality specific independent accident investigation agencies.

This evolution from technical-investigative and sector-specific committees into independent and interdisciplinary based diagnostic instruments for socio-technical systems yields a superior capability to enhance safety, provide a public voice advocating safety, provide transparency in the complexity of systems and contribute to a proper functioning of a civil society. The products of a fully evolved board may serve as input for risk decision making by private and public stakeholders in the management of complex systems during their design and operations. Safety boards may serve as 'problem providers' to other stakeholders in the system. Consequently, fully evolved boards may add to the learning potential of organisations. Moreover, they may serve the integration of safety in a systems safety approach at a socio-technical level.

Four safety Schools of Thought

Safety in modern transportation systems has been an issue for about 150 years. It evolved as a discipline from several different domains and disciplines and has a strong practical bias. Consequently, three 'schools of thought' have been established, which can be categorised as 'Tort Law School', 'Reliability Engineering School' and 'System Safety Engineering School' [4]. In addition a fourth school will be defined as 'System Deficiency and Change' [5].

Each of these schools represent a different pattern of thinking and can be considered as consecutive, representing the societal and scientific safety concepts of their times. They identify specific roles for accident investigation agencies.

Tort Law: The 'Tort Law School' as defined by McIntyre, has a long history and roots in the U.S. railway industry since the end of the 19th century. Out of this development, an engineering design approach emerged, focusing on certification and standardisation of technical designs and products. This development found its counterpart in 'forensic engineering', focusing on technical failure and fact-finding for the benefit of tort and litigation in liability issues concerning accident investigation, mechanical and structural failure of buildings, constructions and products [6]. The concept of failure is central to understand engineering, for engineering design has as its first and foremost objective the obviation of failure [7].

Reliability Engineering: Reliability Engineering became a new engineering school based on the problems of maintenance, repairs and field failures during the second World War. The drive to understand the likelihood of hardware malfunctions and errors, led to the adoption of Probabilistic Risk Assessment in many high-risk industries, among which the process industry and energy supply sector [4].

It was only a natural development that the focus of mechanical reliability engineering expanded to the area of the human factor, predicting human reliability. Cognitive aspects of human error, defining and operationalizing the concept of human failure, expanded from the technical aspects into organisational aspects of systems, examining the complex relation between organisational culture and safety.

Systems Engineering: The Systems Engineering school developed with the dawn of space transportation. This approach focused on accident prevention and was heavily supported by the development of safety standards, specifications and operating instructions. Several accidents in aviation underscored the need to draw a distinction between regulatory compliance for

'certification' and 'safety' when communicating risk to the public [4]. The sociologist Turner defined disaster by its social impact: a significant disruption of existing cultural beliefs and norms about hazards and their impacts. He expanded the technical systems approach into socio-technical systems. As a consequence of expanding scopes, attention should also pay attention to higher order systems levels and post-event consequences dealing with rescue, emergency and crisis management or administrative responsibilities, institutional constraints and policy decision-making and policy management issues. Demarcation lines between investigating major accidents and Parliamentary Inquiries become thin, implicitly restoring the concept of governmental blame.

System deficiency and change: In addition to these three 'schools of thought' a fourth school has emerged during the last decade. Based on the operational experience of Transportation Safety Boards throughout the world, a school of 'safety deficiency and system change' is developing [8]. In this school the concept of independence is crucial, separating the investigative mission and efforts from allocation of blame and vested interests of major stakeholders. This school does not longer focus on 'deviation' from a normative performance or on 'error', but refers to 'system deficiencies'. The focus is on safety critical characteristics of systems in their structure, culture, contents and context with respect to safety critical performance throughout their life cycle [9].

These characteristics can be identified and analysed, based on similarities with other systems, accident and incident data and single case studies. However, such a preventive, encompassing analysis is not always feasible in practice due to the complexity and dynamic nature of transportation systems and the lack of adequate information.

Therefore, a retrospective and independent investigation into systemic incidents, accidents and disasters is indispensable. Such independent investigations may provide a temporary transparency over the actual systems operational performance as a starting point for dealing with inherent deficiencies in such systems.

Independent investigations are considered a right of every citizen and a duty of society. They may put an end to any public concern and can help victims and their families come to terms with their suffering. In addition to learning lessons for the future, independent investigations make our actions transparent and help democracy to function properly [10].

Diverging trends

In retrospect, developments in the various schools of safety thinking have lead to divergence between transportation safety agencies and scientific safety thinking.

Several reasons for such a divergence can be observed:

- Divergence of expertise, experience and knowledge. Historically, accident investigation has been closely connected to technical failure of designs and objects. In their times, scientific notions of reductionism stimulated a strict distinction between research and investigation and fragmentation of scientific disciplines. Forensic engineering, by definition, restricts itself to a supportive expert role for litigation and legal court procedures [6[. The focus of safety boards followed the development of increasing operational complexity of transport systems and technological developments during design and construction, while its counterpart of governance and control developed towards liability and tort law, appointing responsibilities to operators and other stakeholders. The second and third school of thinking elaborated along these lines towards probability, reliability, prevention and societal impact of major events, shifting its focus towards quantification of risk and cost-benefit considerations, crisis and

disaster, emphasising social, behavioural and managerial aspects rather than technological. At present, a need for systems integration and integral safety approaches, facilitates a shift in a reversed direction towards multi-disciplinary co-operation and co-ordination, creating the need for a fourth school of safety thinking and a reconsideration of the role of safety boards. A convergence seems immanent.

- Diversity of focus. In order to investigate complex accidents and draw lessons, safety boards have to focus their attention on the fact-finding phase of occurrences as the start of their investigation process. They have to provide a temporary transparency in complex systems, based on a single-event fact-finding missing. They are not involved in engineering design or certification processes and risk assessment decision making during design, construction or operations. They only have access to detailed knowledge and information once the ultimate test of integrated systems performance occurs during a major event. Their fact-finding strategies have a strong practical basis, in which simplified causation models and lack of structure in underlying factors rather hamper than support a realistic modelling of complex chains of events. Advanced scientific concepts of human performance, based on cognitive psychological principles, are only in their first phases of practical application by accident investigators [11], [12]. Notions of organisational failure and institutional constraints are in their early phases of development, competing each other and lack operational procedures and protocols during investigations. Moreover, analytical models for accident investigation, which are available presently, lack a systems concept, cover a specific range of problem areas or are not codified for more generic applications [13], [14].

Divergence in rationalities

It should be realised that actors involved in the investigations of safety boards may have fundamentally different notions of risk and may apply completely different rationalities [15].

During the conceptual design phase, projects and products are defined by a systemic rationality derived from physics, mechanics, engineering design principles and construction. This phase is linear and confined to specialists. The results of these design decisions are firstly and only exposed to an outsider view and judgement after the detailing phase during testing or operation. Risk perception of operators and users is based on a political and societal rationality. Such rationality is defined by interactions with other actors, negotiating and defining social reality in which operators have to cope with the complex and dynamic operational environment. Decisions made by commissioner and designer have led to a product which can be perceived by its physical appearances without revealing the inherent decisions of the earlier phases. Its operational performance can only be reconstructed by its physical appearance and behaviour as exposed to operators and users. The technology which is applied is therefore ‘to be discovered’ to actors during the operational phase, taking the earlier design decisions as incontestable facts. Characteristics of the design may manifest themselves during the operational phase by incidents, accidents or disaster. Transparency of safety aspects in both rationalities is a crucial issue since safety may be outbalanced and obscured by other interests of a higher order. Such interests may manifest themselves only after an independent investigation into major accidents [10].

Rationality of a designer and engineer focuses on realisation and is reasoning from goal and concept towards function and form. It follows a synthesising and decision oriented line of reasoning. Rationality of an operator and user focuses on perception and knowledge. It follows a line of reasoning from observation, perception, towards structure, function and goal. It is analytic and conclusion oriented.

To understand risks and safety issues two different lines of reasoning are available:

- an ‘inside-out’ vision of commissioners, designers, engineers and other actors which have an oversight of structure and contents of complex systems during their design, development and manufacturing. They are capable of defining complex interactions, couplings and causal relations within the system, risk management, mitigation and control included. They are less capable of dealing with the actual behaviour of the system in its dynamic social environment in terms of risk perception and risk acceptance issues.

- An ‘outside-in’ vision of operators, users, risk bearers, regulators, administrators and other stakeholders which have to cope with the system characteristics in its operational environment. They are capable of dealing with global risk notions and causal relations at an aggregated level, but lack an profound insight into the functioning of complex systems. They may concentrate on perception and acceptance rather than controlling risks.

An ‘inside-out’ vision is likely to define risk in terms of a program of requirements and standards, as a consensus document for the actual design and manufacturing. An ‘ouside-in’ vision is likely to define risk in terms of a defined reality among actors, negotiating risk as a ‘social construct’ to achieve consensus on perception and acceptance between stakeholders. If such a consensus is lacking during events with a high social impact such as disasters, a ‘battleground’ situation may occur, where actors dispute conflicting observations and perceptions.

A second diversion of rationality between accident investigation and scientific research should be taken into account as well. Investigators and researchers both apply a systematic and logic process of reasoning, but these processes have different characteristics.

- the investigative rationality has to deal with the complexity of a reality within yet unknown operating conditions, deals with non-repetitive occurrences and requires a multi-disciplinary involvement of experts. Research rationality deals with the relative simplicity of modelling reality, operates in a controlled environment, is submitted to requirements of repetitiveness, and requires in-depth involvement from one or few scientific disciplines.

- investigations have their starting point in reality, aiming at a fact-finding mission and reconstruction of a time line based sequence of events, focusing on a specific occurrence in its social environment, revealing decision making processes, actions and judgements of participants. Research has its starting point in theoretical expectations of a presumed behaviour of a phenomenon, applying formal logical methods and procedures in order to enable a generalisation of the findings under controlled conditions.

- investigations apply a toolbox of field observations, reconstruction, collection of tactical information on participants to the occurrence and may be supported by specific forensic techniques. Research applies a different type of toolbox, dealing with laboratory controlled experiments, mathematical modelling of phenomenon, controlled data sets, simulation and aims at verifying of falsifying hypotheses.

Consequently, diversity exists between a mission of investigators to establish accident scenarios and system deficiencies based on a robust fact-finding mission and event analysis, and scientists and stakeholders who are interested respectively in specific knowledge aspects, methods or actor-related and discipline-related outcomes of the investigation of the same event.

Convergence

A possible next step in the evolution of safety boards will be towards the role of public safety assessor [16]. Present safety boards already function as gatherers of information across stakeholders and actors. It is a small step into an information dissemination role as well. During the TWA 800 and Swissair 111 disasters, the NTSB and the Canadian TSB took a role of clearinghouse for informing the public and victim's relatives after the disasters. In the near future, safety boards may be seen as safety ombudsmen, the principal advocate for safety and appropriate care of accident victims [17]. They also may expand to the area of rescue and emergency issues, since modern safety boards have a mission in investigating relevant aspects before, during and after the event. TSB's may function as problem providers to other stakeholders in the system, requiring communication skills, risk assessment capabilities and safety management control options. A convergence with other system functions is emerging.

Operating in a multi-actor, multi-stakeholder and multi-rationality environment brings a necessity to reflect on notions and methodologies, which have been are applied in accident investigation. Differences exist across schools of thought, rationalities, sectors and scientific disciplines. If such differences are not recognised properly, accident investigation may take a form of crisis management rather than safety management, implicitly bringing back a notion of blame and liability.

Missions of modern safety boards: The mission of present independent safety board covers four principal objectives;

- determining preventable or mitigable causes of major accidents, disasters and catastrophes in transportation as well as other sectors, irrespective of blame and liability
- identify precursors to potential major events and systemic deficiencies
- increase safety by making acceptable and implementable recommendations
- assure public confidence in safety on a national or sectoral basis.

This mission distinguished TSB's from other investigative authorities such as in-company investigators, governmental accident investigation committees or parliamentary inquiries. The strength of a board for its mission comes from its independence, credibility and ability to address recommendations to any relevant party. Their responses to the board is not only based on a legal mandate of the board to demand timely responses to recommendations but also on the evidence that emerges from its investigations.

Primary working processes: To guarantee a successful mission, five primary working processes of boards have been identified in an international survey of best practices of multi-modal boards in the USA, Canada, Sweden and Finland and a number of single mode boards in the Netherlands.

These five processes of a safety board move the board from the decision to undertake an investigation of one or more accidents or incidents through the analysis of the events into formulations of recommendations to prevent or mitigate future accidents and finally to assessing the effects of those recommendations. Accompanying these actions are ongoing communications with the involved parties [18].

The processes can be characterised in a conceptual model as a benchmark for understanding the evolution of safety boards. The generic model identifies and links the five processes (see figure).

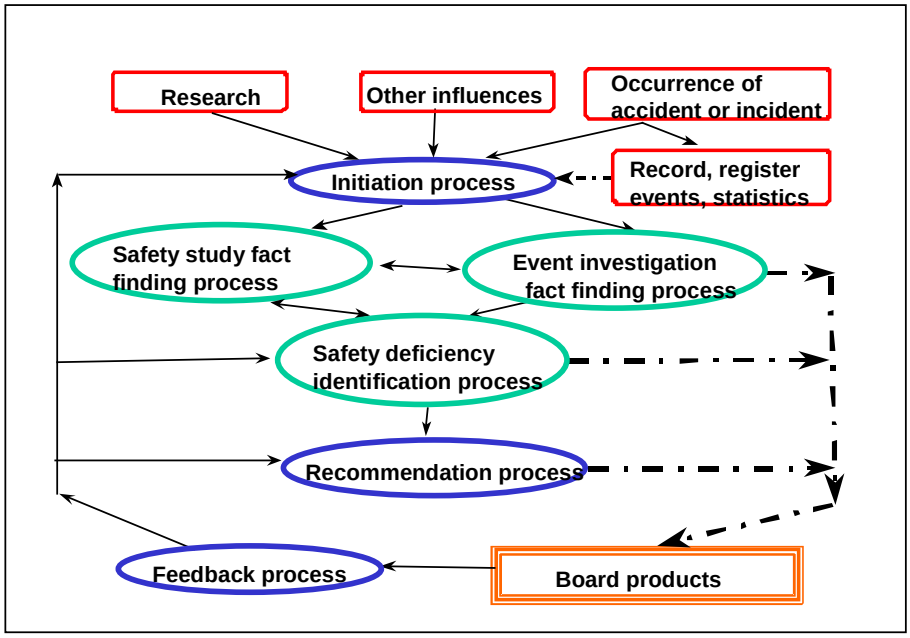


Figure - 1 Five processes define work of a board

These five processes are:

1. an *initiation process* to decide whether to take action or not. A board obtains information about specific transportation accidents and incidents, as well as summary statistical information on transportation conditions and events and the results of research relevant to transportation safety. In the case of specific events, the board has a mechanism that helps it decide which events merit an intensive investigation.
2. A *fact-finding process* to assemble all relevant data bearing on an event and to determine findings about the main factors contributing to the event or general situation. There are three forms that the fact-finding may take: a reactive *event investigation* of an accident or incident constituting the majority of most boards efforts, a *retrospective safety study* to attempt to determine the factors associated with and preceding events or a *pro-active safety study* in which the board plans a research study that includes primary data collection of events as they occur.
3. A *safety deficiency identification process* that takes the facts at hand derived from single events or from safety studies, and determines systematic threats to transport safety. The safety deficiency identification process can use modern scientific tools such as pattern recognition, multivariate regression, functional decomposition, task analysis, dynamic systems modelling or can be based on operational experience or a combination of these two.
4. A *recommendation process* that formulates effective steps to prevent or mitigate the harms of accidents and incidents. These steps should be also economically and politically assessed in

order to comply with their social acceptance and sustainable effects. The recommendation process may include considerations of how proposed actions might be implemented.

5. A *feedback process* that maintains contacts between the work of the board and the external public world. A central feature of this feedback process is a systematic monitoring of the recommendations of the board, both in terms of the actions taken in response to the recommendations and the effects of these actions on transportation safety.

Interfacing issues: In order to establish a working relation across investigators, researchers and stakeholders, a conceptual model of the accident investigation process has been established and the need for an investigation methodology has been recognised.

Such processes and methodologies should facilitate the required convergence between safety boards and their operating environment in view of their new missions and independent position.

A probe into the nature of the investigation process reveals several characteristics, which control the steps of the process and their interfaces:

1. the initiation process.
   This phase realises the transition *from symptom to syndrome*. In contrast with statistical and epidemiological analyses, which focus on isolated or specific contributing factors in accident causation, pattern recognition and trend analysis may reveal context specific combinations of factors that provide necessary and sufficient conditions for the causation of the event. This medical model refers to causation and context instead of correlation and may be applied to identify 'investigable' accidents by their type or classes. The transition from symptom to syndrome facilitates selectivity in the investigation process and may lead to safety studies of specific events or a single event investigation.

2. the investigation process.
   This phase defines the transition *from fact to factor*. Based on a fact-finding mission, the investigation derives a set of events in their time-dependent sequence, which together provide a satisfactory explanation of the occurrence. A wide variety of causation models and notions has been applied in practice, referring to 'underlying causes', 'contributing factors', 'primary' and 'secondary causes' etceteras. Taking into account the nature of the various schools of safety thinking, a categorisation of causality may be derived in four consecutive categories:

- deterministic causality. Related to the first school, this form of causality refers to a static relation between system characteristics. This causality has been applied in particular in engineering design, leading to design principles such as fail safe, crash worthiness, damage tolerance, containment, zoning, etc. This category relies on insights in failure modes and performance envelope parameters.

- probabilistic causality. This category has its roots in reliability engineering developments, referring to the probability of occurrences, related to RAMS principles and data analysis of past performance of similar systems. This category relies on sufficient data of similar nature, expert judgement and other sophisticated risk estimate tools.

- intentional causality. In view of the third school, a third category of causality was added to the scope of the investigation, dealing with motives for decision making. This causality distinguishes the investigation process from a judicial inquiry due to the fact that the focus in no longer on intentional and possible criminal behaviour, focusing on individual motives,

means and opportunities to commit an act. However, normative notions are still frequently applied referring to human error at the operator level.

- situational causality. This category refers to the complexity and dynamic behaviour of systems under specific operational conditions, due to which accidents and incidents may occur. Unanticipated coincidences may occur due to a mismatch in synchronisation, commonly referred to as 'wrong time, wrong place' type of events.

  These categories of causation however, refer to the phenomenon of 'cause' which still has an implicit normative notion of blame or liability. Consequently, it puts a 'burden of proof' on the agency that establishes the 'causes' of accidents, referring to similar mechanisms in forensic investigations and still bears similarities with judicial procedures. The concept of 'deviation' from an implicit normative standard is still present.

  Therefore, to conduct independent investigations, a scientific method should replace the implicit judicial procedures and protocols. Due to the nature of the investigation process, a promising approach might be provided by a case-study methodology as defined by Yin [19]. The approach will serve as a starting point for the investigation process by fact-finding on the accident site of a single occurrence. Since such an approach will not be without theoretical framework, the data collection strategy will be provided by the systems concept, supported by forensic and analytic techniques. Theoretical assumptions are provided by a first definition of possible accident sequences, elaborated by further analysis, validation of data collection and additional scientific methods such as pattern matching, explanation building and time-line analysis. A satisfactory explanation of the occurrence should be the result, represented by one or a few accident scenarios.

3. the process of safety deficiency identification.
   This phase defines the transition *from deviation to deficiency*. In order to structure the identification of systemic deficiencies from the previous phase, systems modelling has to take place. Such a modelling should facilitate a satisfactory explanation of the overall systems safety performance and indicate where and how system characteristics have contributed to the safety deficiencies. Unfortunately, a comprehensive system modelling is not yet readily available for investigation purposes. In practice, a system modelling depends on the available models within various scientific disciplines. Theoretical models are developing, taking into account a systems level hierarchy, life cycle approaches, strategic decision making structures or engineering design processes [20], [21]. Remedies for enhancing system deficiencies are found in technological engineering design principles on a conceptual level and in specific intervention strategies, such as defence barriers, based on the model developed by Reason.

4. Drawing up recommendations.
   This phase defines the transition *from identifying explanatory variables to control variables.* Having a satisfactory insight into the origins of system deficiencies does not imply a control over their actual conduct. In any complex and dynamic system, constraints and conditions may be present, obstructing safety enhancement measures of any nature. Such variables may be of a natural origin or may be defined by institutional constraints or long term synchronisation problems across system life cycle boundaries. Technological or conceptual innovation in the systems structure, organisational culture or primary processes may be required in order to change the safety performance of the system. Historically, recommendations have primarily focused on elimination of causal factors, rather than on improving the learning potential of a system at higher organisational levels. A proper control over 'underlying factors' or 'secondary causes' should be related insights into the dynamic

behaviour of the system regarding risk management strategies, regulatory and institutional levels in the system and eventually, societal values and norms.

5. Monitoring and feedback.
   This phase defines the transition *from control options to risk assessment and risk communication* in order to achieve cost-efficient and sustainable societal support for safety enhancement measures. It is debated whether or to which extend this transition is an intrinsic part of the investigation process itself in view of the required independence of the investigations. It can be debated that an objective diagnosis of the occurrence and identification of system deficiencies does not include involvement in the actual implementation of the recommendations for systemic improvements and risk mitigation. However, the shift in focus and mission expansion indicates a trend towards further involvement of safety boards in this process. Such an involvement might require new qualifications and tools to accommodate such involvement in terms of risk assessment techniques, risk communication and expanding focus of the investigator towards all life cycle phases and operational processes of a system. Developing an investigation methodology may become necessary in the future to compensate for the accumulation of operational experience and knowledge of major players in the investigations. Changes in engineering design methodology with respect to collaborative and knowledge based engineering may put additional demands on the investigative skills.

   At present the Canadian Transportation Safety Board explicitly applies the full scope of the 5 principal processes [22]. The goal of its ISIM methodology (Integrated Safety Investigation Methodology) is to strengthen the integration of the investigation, safety deficiency analysis and communication process. The methodology aims at helping investigators to identify risks in the transportation system by co-ordinating all aspects of the investigation process. The method emphasises the concept of iterative investigations, providing a way to maintain an overall understanding of an occurrence while on-going data collection, analysis, and communication are carried out. Consequently, the concept has abandoned the notion of a final report, discussing findings and recommendations in public.

Discussion and conclusions

A fundamental reason to introduce independent accident investigation was that parties involved began to realize that criminal law inquiries focus on allocating blame. To learn lessons for the future and to take steps to prevent similar accidents, it was essential to identify the causes of these accidents. Another type of investigation was thus needed. From a judicial point of view however, investigation methodology is restricted as the more useful tool for criminal intelligence analysis. It has strong ties with conventional ‘forensic engineering’ methodologies applied to determine liability for structural failure in engineering design. A clear distinction is made between various forms of logic reasoning, by applying either the notions of ‘investigation’ or ‘research’. ‘Research’ based methodologies have been considered less useful for a fact-finding phase of investigations, since their inference do not go beyond the premises of their scientific discipline, not arriving at any new causes, conclusions or recommendations. In addition, the scope of criminal inquiries was restricted to discovering the direct cause of an accident and to identify an unacceptable deviation from a normative standard, not the underlying causes or systemic deficiencies. This was aggravated by the fact that suspects were permitted to withhold information not to incriminate themselves. Conventional accident investigation methodologies therefore, tended to focus on cause and not on prevention.

In adapting to changes in the working environment of TSB's, not only the products and methodologies of the boards are changing, their mission, role and position to other stakeholders are changing as well. TSB's might be assessed along lines of a product development cycle themselves. From a product life-cycle point of view, TSB's enter a next phase in their existence. Starting as a technical committee, focusing on causal and forensic aspects with a pre-event focus, they gained an influential and credible position within several transportation modes. In a second step, their scope expanded towards non-technical aspects and higher systems levels, such as human error, organizational failure, gaining independence from allocation of blame and governmental influence. In a third phase, external influences were incorporated in the TSB working processes such as rescue and emergency aspects, victim care and family assistance. In a next phase, TSB's might develop new mission elements, participating in a knowledge network, dealing with risk assessment approaches, communication with stakeholders and providing safety control options for stakeholders during design and operation of complex systems.

It may be concluded that independent Transport Safety Boards represent a distinct school of thought in accident investigation. Historically, they have strong relations with engineering design and identifying failure in technical systems. Transportation Safety Boards however are evolving towards a socio-technical systems approach. Several methodological issues are yet to be resolved to guarantee their independence, credibility and reputation as a qualified agency. Historically, the role of fact-finding and accident reconstruction has firmly been established in the relation to engineering design and operations in transportation. New sectors and scientific disciplines have emerged and working relations are established with other high-tech industrial sectors.

TSB's need to develop their own methodology to comply with the need to link the processes of fact-finding, establishing system deficiencies to the process of drawing up recommendations and implementing systemic changes. It may be necessary to combine these processes in an appropriate form, despite the fact that fundamental differences exist between risk notions, rationalities across actors, stakeholders, investigators and researchers and their objectives in an accident investigation process. It also clarifies the need for the Transport Safety Boards community to participate in an information infrastructure because TSB's will not be able to cover all required expertise on an in-house basis. It may be stated that in addition to a formal and functional independence, TSB's may also need to develop and maintain methodological independence.

References

[1] De Kroes and Stoop 1992
First World Congress on Safety of Transportation. Delft, 26-27 November 1992. Delft University Press

[2] Hengst, Smit and Stoop 1998
Second World Congress on Safety of Transportation. Imbalance between Growth and Safety? Delft, 18-20 February 1998. Delft University Press.

[3] ETSC 2001
Transport accident and incident investigation in the European Union. European Transport Safety Council. ISBN 90-76024-10-3. Brussel 2001

[4] McIntyre 2000
Patterns in Safety Thinking. A literature guide to air transportation safety. Ashgate Publishing Ltd

[5] Stoop 2002
Accident investigations: trends, paradoxes and opportunities. International Journal of Emergency Management . Vol 1, No 2, 2002, pp 170-182

[6] Carper 1989
Forensic Engineering. CRC Press, First Edition, 1989

[7] Petroski 1992
To engineer is human. The role of failure in successful design. Vintage Books

[8] Johnson 1999
Organization and activities of the TSB. Lessons from Independent Accident Investigation in Canada. Presentation to the Transportation Safety Conference Kansai University, 22 July 1999

[9] Stoop 1990
Safety and the design process. Doctoral Thesis Delft University of Technology. Delft University Press.

[10] Van Vollenhoven 2002
Independent Accident Investigation: Every Citizen's Right, Society's Duty. Chairman Dutch Transportation Safety Board, Chairman International Transport Safety Association, (Founding) Board Member European Transport Safety Council. The Hague, the Netherlands

[11] Dekker 2002
The field guide to Human Error Investigations. Cranfield University press, Ashgate

[12] Strauch 2002
Investigating human error: incidents, accidents and complex systems. Ashgate

[13] Rimson and Benner 1996
Mishaps investigations: Tools for Evaluating the Quality of System Safety Program Performance. In: Proceedings 14th International System Safety Conference, august 12-17, Albuquerque, New Mexico. pp 1C2-1 – 1C2-9

[14] Sklet 2002
Methods for accident investigation. Norwegian University of Science and Technology. Dept of Production and Quality Engineering. Trondheim, Norway

[15] Stoop 1996
Risicobeheersing bij technisch-complexe projecten. In: Grote projecten, besluitvorming & management. Editors: De Bruijn, De Jong, Kortsen and Van Zanten. Samson H.D. Tjeenk Willink

[16] Kahan, Frinking and De Vries 2001
Structure of a Board to Independent Investigate Real and Possible Threats to Safety. RAND Europe, May 2001

[17] Hovden 2001
Regulations and risk control in a vulnerable society: points at issue. International Conference on Emergency management. TIEMS 2001, June 19th-22nd, Oslo

[18] Kahan 1998
Safety board methodology. In: proceedings of Second World Congress on Safety of Transportation. 18-20 February 1998, Delft University of Technology. Editors S. Hengst, K. Smit and J.A. Stoop

[19] Yin 1994
Case Study Research, Design and Methods. Second Edition Applied Social research Methods Series. Volume 5. SAGE Publications

[20] Evers, Bovy, De Kroes, Sommerhalder en Thissen 1994
Transport, infrastructuur en logistiek: een proeve van een integrerend onderzoeksprogramma. In Dutch. TRAIL Onderzoeksschool, Delft University of Technology, February 1994.

[21] Rasmussen and Svedung 2000
Proactive Risk Management in a Dynamic Society. Swedish Rescue Service Agency. Karlstad, Sweden

[22] Ayeko 2002
ISIM; investigating for risk mitigation. Workshop on the Investigation and Reporting of Incidents and Accidents. 17-20 July 2002, University of Glasgow

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01- 09 - 2003	Conference Publication	

4. TITLE AND SUBTITLE

Second Workshop on the Investigation and Reporting of Incidents and Accidents, IRIA 2003

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Kelly J. Hayhurst and C. Michael Holloway, Compilers

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

23-765-30-10

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

NASA Langley Research Center
Hampton, VA 23681-2199

8. PERFORMING ORGANIZATION REPORT NUMBER

L-18324

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSOR/MONITOR'S ACRONYM(S)

NASA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

NASA/CP-2003-212642

12. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified - Unlimited
Subject Category 03
Availability: NASA CASI (301) 621-0390 Distribution: Standard

13. SUPPLEMENTARY NOTES

An electronic version can be found at http://techreports.larc.nasa.gov/ltrs/ or http://ntrs.nasa.gov

14. ABSTRACT

This publication consists of papers presented at the Second Workshop on the Investigation and Reporting of Incidents and Accidents, IRIA 2003, sponsored by NASA Langley Research Center and the University of Virginia.

15. SUBJECT TERMS

Accidents; Mishaps; Incidents; Causal analysis; Software; Investigation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			STI Help Desk (email: help@sti.nasa.gov)
					19b. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	254	(301) 621-0390